1-1-1978

# Representation and processing of logical relations in meaningful text: linear orderings and set inclusions.

Susan Bennett Sefkow
*University of Massachusetts Amherst*

# REPRESENTATION AND PROCESSING OF LOGICAL RELATIONS IN MEANINGFUL TEXT: LINEAR ORDERINGS AND SET INCLUSIONS

A Dissertation Presented

By

SUSAN BENNETT SEFKOW

Submitted to the Graduate School of the University of Massachusetts in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 1978

Psychology Department

REPRESENTATION AND PROCESSING OF LOGICAL RELATIONS IN MEANINGFUL TEXT: LINEAR ORDERINGS AND SET INCLUSIONS

A Dissertation Presented

By

SUSAN BENNETT SEFKOW

Approved ~~as~~ to style and content by:

Charles E. Clifton, Jr., Chairperson of Committee

James I. Chumbley, Member

Grace J. Craig, Member

Marvin Daehler, Member

Jerome L. Myers, Member

Bonnie Strickland

Bonnie Strickland, Department Head
Psychology Department

ABSTRACT

Representation and Processing of Logical
Relations in Meaningful Text:
Linear Orderings and Set Inclusions

September 1978

Susan Bennett Sefkow, B.A., Yale University
M.S., Ph.D., University of Massachusetts
Directed by: Professor Charles E. Clifton, Jr.

The present experiment was conducted to examine conflicting findings reported in comparing the processing of two very similar types of linguistic material: set inclusions and linear orderings. Consistently Ss perform better on remote than adjacent linear relationships. A very different pattern is obtained with set inclusions. Memory for true adjacent set inclusion relations exceeds remote relations while just the opposite holds true for false items. Potts (1976) has argued that both set inclusions and linear orderings share the same form of integrated memorial representation, but that Ss tend to respond erroneously to set inclusion relations on the basis of similarity rather than actually evaluating the relation. This shared representation hypothesis was tested within a paradigm which also addressed another issue.

Sefkow (1976) recently demonstrated that the processing initiated by a correct attempt to answer a question about

information available only in memory could substantially facilitate retention of that material. This backward review effect was replicated and the nature of the process responsible for the effect was examined.

Ss listened to five equivalently structured prose passages based on either set inclusions or linear orderings. Presentation order of the four adjacent relations within each passage was varied to manipulate the ease by which the relations could be integrated. Immediately after each passage Ss were asked to verify either a true inference drawn from the passage or a false statement. Subsequent free recall and recognition data were collected under intentional learning instructions. Correctly verifying a true inference involving two to four mediating set inclusion relations resulted in later enhanced recall of those same relations, independent of integrability. Only when initial integration of the linear orderings was hampered and Ss were presumably forced to rely on memory for the presented relations were similar recall patterns obtained with linear orderings. These data contradict Potts' proposal that the two relation types are similarly represented in memory. It was suggested that the unidimensional character of linear orderings makes the integration stategy apparent. The strategy may also be necessary in order to avoid the individual linear relation's susceptibility to interference as was evidenced by the poor overall performance obtained when

integration was hampered. On the other hand, it was suggested that set inclusions are subject to an encoding strategy which focuses on the elements' shared semantic attributes implicitly specified by set inclusion relations. The the presented relations may tend to be stored as multi-dimensional units relatively independent of one another.

Table of Contents

List of Tables

List of Figures

## Introduction

An ever increasing literature suggests that people, as active processors of information, subject linguistic inputs to an abstractive, constructive encoding (e.g., Barclay, 1973; Bartlett, 1932; Bransford, Barclay, & Franks, 1972; Bransford & Franks, 1971, 1972; Bransford & Johnson, 1973; Cofer, 1973; Jenkins, 1974). Comprehension does not involve the interpretation and storage of single sentences. Rather, information is integrated within and between sentences and is further augmented by the person's world knowledge. How can such a hybrid be characterized? Potts and his associates (e.g., Potts, 1972, 1974a,b; Scholz & Potts, 1974) have extensively investigated this constructive process using one type of verbal material.

Potts' (1972) original work began as an attempt to determine whether or not people actually store deducible information in memory. Quillian (1969) had proposed that, in the interest of "cognitive economy", people store only necessary information and deduce the remaining whenever it is needed. Other researchers (e.g., Anderson & Bower, 1973) argued that, on the contrary, people may indeed store the deducible information. To decide between these hypotheses, Potts had subjects study paragraphs describing arbitrary linear orderings. Each consisted of four nouns related by a comparative adjective which can generally be represented by

A>B>C>D. Subjects studied only the three adjacent relations (A>B, B>C, C>D) sufficient to describe the ordering. Transitivity allows the deducement of three remote relations (A>C, A>D, B>D). After studying each paragraph, subjects were given a true-false recognition test. All six true relations, as well as six false relations formed by reversing the correct pairings, were employed. Potts found that subjects were consistently more accurate in responding to remote pairs than adjacents, despite the fact that they had not been presented. This result held for both true and false pairings.

Subsequent studies replicated the effect with longer orderings and also found that reaction times paralleled the error data. Subjects were faster to verify relationships between two items, the further apart they were in the ordering. This distance or "stepwise" effect was complicated somewhat by "end anchor" effects. Performance was generally enhanced by the presence of an end term in a test item (e.g., A or D in the four term ordering) and this was confounded with distance. By using longer orderings (Scholz & Potts, 1974) the two effects were separated and both were found to determine performance. The ease with which an ordering is constructed is significantly affected by another variable: the original presentation order of the adjacent pairs (Smith & Foos, 1975; Smith, Foos, Sabol, & Mynatt, 1976). When new pairs add an optimal amount of information to memory (one new item) and do not necessitate reorganization of the developing

linear ordering, performance is at its best. These data are offered as strong support for the constructivist position. Clearly, subjects make and store inferences while studying.

This "symbolic distance effect" (SDE), so named by Moyer & Bayer (1976), has been found to be a widespread phenomenon. The greater the psychological distance between two items which can be ordered along some dimension, the faster people can compare their magnitude. Besides occurring with arbitrary orderings acquired in the lab, the effect extends to judgments made about digits, (e.g., Moyer & Landover, 1967; Parkman, 1971; Buckley & Gillman, 1974), letters (e.g., Parkman, 1971), object sizes (e.g., Moyer, 1973; Paivio, 1975; McKinley, 1976), prototypicality of category instances (e.g., Rips, Shoben & Smith, 1973), and other natural language concepts such as time, temperature, etc. (Holyoak & Walker, 1976). The effect has been demonstrated with children, as well as adults (e.g., Riley & Trabasso, 1974; Trabasso, Riley & Wilson 1975) and parallels an entire literature studying perceptual comparison processes (cf. Moyer & Bayer, 1976).

While the existence of the SDE has certainly been established, no single model has been proposed to adequately explain all the data. A wide variety of models have been developed which differ along a number of dimensions (cf. McKinley, 1976). For example, the Potts' model (1974) attributes the distance effect to the discriminability of the items placed on an imaginary spatial array or rating scale. Humphreys (1975)

has suggested that subjects learn the frequency with which items occur as the greater (or lesser) member of a pair and use this as an index of comparison. Others (e.g., Moyer & Bayer, 1976; Holyoak & Walker, 1976) argue that the distance effect is the result of a comparison of stored absolute magnitude information. To account for the effects obtained using arbitrary, experimentally acquired information, at least one point must be granted. That is, a successful model must assume that the interrelationships of all items in the ordering are fully specified in memory prior to testing. This must be at least to the degree that the stored information allows the direct assessment of relationships without the necessity of reevaluating the presented pairings.

Set inclusions (e.g., All X are Y.) are another type of relation very similar to linear orderings in that they share such properties as transitivity and irreflexivity. Sets of chained relations of both types can be elaborated or seriated into a directed linear array. It is therefore surprising to find that set inclusion recognition and recall accuracy data, collected under conditions similar to those described above, have been in direct conflict with the SDE (Frase, 1969, 1970). Memory for adjacent relations exceeded remote relations and appeared to be a decreasing function of distance. One purpose of this research is to take a closer look at this discrepancy and the implications it has for the nature of linguistic memory representations. Relevant evidence will first be reviewed, then questions for investigation will be formulated.

The apparent contradiction was first investigated by Griggs (1976) and independently replicated by Carroll & Kammann (1977) in order to see whether set inclusions would be processed in the same way as linear orderings under comparable conditions. Both investigations employed the basic Potts paradigm and recorded accuracy data. Carroll & Kammann also measured response latencies. Despite their similarity, the two types of relations yielded quite different patterns of results. Overall linear ordering performance was superior to that of set inclusions. While the usual distance effect was manifested by the linear ordering data, a truth by distance interaction was obtained with set inclusions. Accuracy decreased as a function of distance for true items and increased for false. Likewise, reaction times increased with distance for true items and decreased for false. It was concluded that the effect was due to initial processing of the relations and not to a memory deficit because the results were replicated when subjects were allowed to refer to the paragraphs while responding to test items. Griggs suggested that this pattern was the result of two illogical processes specific to set inclusion. First, subjects incorrectly assume symmetry of relations. Given the sentence "All A are B.", "All B are A." tends to be judged true. This is exemplified by the poor performance on adjacent false items. Such invalid conversions have also been found in the syllogism

literature (e.g., Ceraso & Provitera, 1971; Johnson, 1972). Second, subjects are unwilling to assume transitivity and are less so as the number of relations involved increases. This tendency to respond "false" as a function of distance is characterized by an improvement on false items and decrement on true items. These two "processes" accurately describe the data but they hardly provide any real explanation of the phenomena.

In one condition, Griggs instructed subjects as to the validity and invalidity of symmetric and transitive inferences. He also gave them a practice paragraph with all the set relations fully and explicitly elaborated. This was sufficient to eliminate differences between the linear orderings and set inclusions, thus the SDE was obtained. On the other hand, Carroll & Kammann found that general instructions explaining transitive inferences and practice with feedback did not influence the "illogical" processing of set inclusions. Apparently, it is possible for subjects to accurately deduce the remote set inclusion relations, but this does not appear to happen as a natural consequence of processing as in the case of linear orderings.

The results of these studies imply that the two set theoretic relations are initially processed or encoded quite differently. No good theoretical account as to why this is so has yet been offered. Carroll & Kammann acknowledged that, at least within this paradigm, people do not use the

logical processing strategies one would expect. One possible implication is that set inclusions are not readily subject to a constructive or schematic encoding. Alternatively, the set inclusions may be schematized but not as expected based on a "logical" analysis. They further suggested that the problem might lie with the artificial materials used in these investigations. When more meaningful set inclusions are evaluated, knowledge drawn from semantic memory can readily be used to determine such things as asymmetry. Admittedly, the role of semantic memory is attenuated in such studies, but exactly what this role might be is yet to be determined.

Potts (1976) pointed out the occurrence of both patterns of results in the semantic memory literature (e.g., Collins & Quillian, 1969, using set inclusions and Moyer, 1973, using linear orderings), indicating that the findings are not simply due to the artificiality of the materials employed. Furthermore, he suggested that set inclusions and linear orderings may share a common schematization. Differences in processing or retrieval strategies could then be responsible for the contradictory distance effects.

Potts proposed that the set inclusion data might best be accounted for by a process reflecting the tendency of subjects to make such decisions on the basis of similarity. A semantic memory theory, the Smith, Shoben, & Rips (1974)

feature comparison model, was offered as a possible characterization of the process. When a subject is asked to judge whether an A is a B, he makes an initial assessment as to the degree to which A and B share similar semantic features. If they are very similar, the subject responds true; or if they are very dissimilar, the subject responds false. If they are of intermediate similarity, second stage processing is necessary. Here the actual relationship is assessed and it is determined whether the essential features of B, which determine category membership, are shared by A. This model accounts nicely for the interaction of truth value with remoteness obtained with set inclusions. In addition, such a model would account for the two types of logical errors described by Griggs (1976). First, similarity relations are symmetric: if A is similar to B, then B is similar to A. Second, similarity relations are nontransitive: if A is similar B and B is similar to C, it is not valid to conclude that A is similar to C.

To investigate this hypothesis, Potts compared performance on artificial linear orderings, set inclusions, and similarity relations. Individual, as well as group, data supported the notion that there is a strong tendency for set inclusion subjects to respond on the basis of similarity (or some correlate) but that this is not totally pervasive. More importantly, the reaction times of those subjects who did not make logical errors replicated the usual linear

ordering effects. In addition, Potts tested subjects who again studied either set inclusions or linear orderings, but here the terms in the relations were nonsense syllables. For a reason yet to be ascertained, accuracy levels were extremely high, the tendency to make logical errors disappeared, and the latencies of both groups demonstrated the SDE. This result certainly puts to rest the Carroll & Kammann suggestion that the SDE is not obtained with experimentally induced set inclusions because of the artificiality of the materials.

Based on the result that errorless set inclusion performance is accompanied by the typical linear ordering latency patterns, Potts argued that both types of relations are similarly represented in memory. The set inclusion errors occur because subjects tend to respond on the basis of similarity rather than fully evaluating the relations.

In summary, the data suggest that two types of linguistic material which share structural and logical properties are processed in very different ways. Linear orderings are easily assimilated into a memory representation which specifies at some level the interrelationships of all the elements. The deducible information is not so readily accessible for set inclusions. Errors occur even in nonmemory tasks, implying that the relations differ at least at the level of initial processing. When such errors are not made, performance appears to parallel that of linear

orderings. While the source of these errors has yet to be identified, several hypotheses can be formulated as to the nature of the memorial representation of set inclusions. For example, Potts has suggested that linear orderings and set inclusions share the same form of integrated representation but that the latter is susceptible to inappropriate or less than efficient processing strategies at retrieval. Another viable possibility is that the processing errors, whatever their source, either inhibit the formation or reflect the lack of an integrated schema. Thus, set inclusions may tend to be stored as relatively independent propositional units. These hypotheses provide the basis for several differential predictions to be developed within the context of a very different but relevant paradigm which will be considered next.

A considerable body of research exists which is concerned with the effects of asking people questions about textual materials shortly after exposure to it (cf. Anderson & Biddle, 1976). The concern is with performance on the questions but more importantly with the consequences of those questions for learning and memory. The practical implications for education are apparent and explain why the bulk of such research has been done by educators and educational psychologists. However, the basic paradigm can also be employed to investigate more global issues of human information processing. The present research is

proposed with that intention.

The most interesting finding of the adjunct question literature is that when subjects are required to respond to questions shortly after exposure to the relevant text, their memory for that material is enhanced as measured by a later criterion test. Experimental subjects outperform controls not only on the same questions repeated as criterion items but also on new or incidental items. The original explanation for the latter effect (Rothkopf, 1966) was that subjects responded to questions embedded in prose by modifying their processing of subsequent material in order to maximize their performance on succeeding questions.

Watts and Anderson (1971) suggested an alternative explanation. They proposed that the memory search initiated by the adjunct questions might somehow be responsible for enhancing later recall of not only question specific but also nonspecific information. Several subsequent studies addressed this issue (e.g., McGaw & Grotelueschen, 1972; Rothkopf & Billington, 1974) and, indeed, found evidence to support such a backward review effect, though it appeared to be quite small. These studies were performed without reference to the structure of the materials beyond identifying question relevant and irrelevant information. Nor was there any consideration of the processing requirements of the questions or of the nature of the structure(s) in memory to which the questions were directed.

If any clear understanding of the phenomenon is to be had, these issues cannot be avoided.

A recent study by Sefkow (1976) began to investigate the backward review effect within such a framework. The research presented here is intended to continue this endeavor, as well as contribute to an understanding of the structure of linguistic information in memory and the processes which operate on it. The Sefkow study will first be reviewed and then proposals for additional research will be formulated.

Five prose passages very similar to those used in the studies described earlier were employed by Sefkow. Each passage was made up of four set inclusion relationships ($A \subset B$, $B \subset C$, $C \subset D$, $D \subset E$) from which six could be inferred ($A \subset C$, $B \subset D$, $C \subset E$, $A \subset D$, $B \subset E$, $A \subset E$). One of the passages is presented in Table 1. Immediately after listening to each passage, the subjects were asked to verify either one true inference drawn from the prior passage or a false statement. Accurate performance on the true probes required the integration of two adjacent relations (see Table 1). The false items were constructed so that they could be rejected on the basis of unfamiliarity, therefore minimizing the possibility of any meaningful review of the passages. Thus the subjects listened to all five passages, each paired with a different probe ($A \subset C$, $B \subset D$, $C \subset E$, or one of two false statements) subject to appropriate

Table 1

Passage and Probe Examples from Sefkow (1976).

---

Recently, some new facts have been discovered about our universe. According to all present evidence,
$D \subset E \rightarrow$ all planets in Galaxy IV are capable of supporting life.
There is also general agreement on the fact that
$C \subset D \rightarrow$ all the "blue" planets are in Galaxy IV. The
$B \subset C \rightarrow$ only interesting planets are these "blue" planets; however, later discoveries may change this judgment. Scientists consider
$A \subset B \rightarrow$ all known planets within fifteen light years distance to be interesting. It is hoped that new discoveries will be made in the near future which will extend this knowledge.

Possible True Probes:

$A \subset B$ & $B \subset C \Rightarrow A \subset C$: All known planets within 15 light years distance are "blue" planets.

$B \subset C$ & $C \subset D \Rightarrow B \subset D$: All the interesting planets are in Galaxy IV.

$C \subset D$ & $D \subset E \Rightarrow C \subset E$: All the "blue" planets are capable of supporting life.

---

counterbalancing. Subsequent free recall data were collected under both incidental and intentional learning conditions. Evidence for a substantial review effect was obtained: recall of true-probed passages exceeded that of false-probed passages. Specifically, the recall advantage was concentrated on the adjacent relations whose integration or evaluation was necessary for the true probes' verification. Mean recall of these relations ranged up to 24.8% greater than relations from both true- and false-probed passages which were irrelevant to the probes.

One additional finding ruled out the possibility that the elevated recall was due to some cueing or retrieval process: when subjects listened to the passages and were then given the true probes exclusively as retrieval cues at the time of recall, the effect disappeared. When presented with a valid inference (e.g., A $\subset$ C), Ss did not tend to regenerate the constituent relations (A $\subset$ B, B $\subset$ C).

It was suggested that a strengthening or integration of the memory traces at the time of the probes was responsible for the elevated recall. That is, the probes may direct attention to the relevant relations while they are still available in memory, at which time they are rehearsed and thus more readily recalled. On the other hand, success on the verification task requires subjects to evaluate and, in turn, integrate the two relationships to

draw the appropriate inference. It may be that this integrative or more elaborate reencoding (cf., Craik & Tulving, 1975) of the stimuli is responsible for the enhanced recall. Note that these hypotheses are not necessarily mutually exclusive. Risking gross simplification, one question of interest is whether a significant change in the structure of the memory representation occurs as a result of an inferential probe or is the existing structure merely strengthened? This brings us back to more basic questions which were considered earlier: what is the nature of this memory representation to which we refer? Do set inclusions tend to be stored as individual relations or are the implicit interrelationships readily recognized and incorporated into some more global, integrated store as is apparent for linear orderings? The distinction to be made here between an integrated and a propositional store is general. An integrated store is one which allows the direct assessment of relationships between all elements in an ordering without the need to reevaluate any mediating relations. A proposition based store is one which maintains the integrity of the individual presented relations. Inferences between items in an ordering are not represented directly but must be drawn by consulting the mediating links.

The research reported here addresses a number of issues. The generality of the backward review effect is tested using new set inclusion materials, as well as linear

orderings. Prior investigations of these two relations have focused almost exclusively on recognition accuracy and latencies. The probe manipulation and attendant free recall data collected here should provide a valuable alternative approach. Comparisons between the two types of relations may reveal effects which will help to explicate both the process responsible for the review effect and the nature of the memory representations involved.

## General Procedure and Rationale

In the present research subjects were required to respond to inferential probes based on memory for either set inclusions or linear orderings. An intentional free recall task followed and recognition data were collected where possible. The original Sefkow (1976) study was replicated first with both types of materials to determine whether similar patterns of facilitated recall are obtained. Probe error rates were also compared. These data serve as a frame of reference to which the following conditions are contrasted. Condition 2 investigated the importance of order of presentation of the relations within the paragraphs to probe error rates and patterns of facilitation in the recall data. Finally, in a third condition, true probes were used which span two and three, rather than just one, intervening item. Again, the probes' error rate and effects upon free recall were evaluated.

## Backward Review Effects: Set Inclusions versus Linear Orderings

Probes. As in the Griggs (1976), Carroll & Kammann (1977), and Potts (1976) studies, a substantial error rate was obtained by Sefkow (1976) when subjects were asked to verify transitive inferences involving set inclusions. Errors on the probes which required a two-link transitive inference occurred on approximately 28% of the trials.

Subsequent to the free recall task, Sefkow collected recognition accuracy data under verification instructions. Here too the data replicated the previous finding that as distance between items in the true test pair increased, accuracy decreased. Sefkow had hypothesized that the probe errors were due to a deficit in memory for the presented relations. The illogical processing hypothesis of Griggs provides a viable alternative characterization. Should the errors be due to faulty logic specific to set inclusions, then a replication of the study using linear orderings should greatly reduce the number of such errors. On the other hand, two types of memory deficits are possible. First, if linear orderings yield lower error rates it may be because of differences in the difficulty of encoding. Should both recall and recognition of the presented information be equivalent for the two relations, then this type of deficit can be rejected. Second, the errors may be due to a memory failure occurring after input. In this case, probe errors for the two types of relations based on identically constructed paragraphs should be similar. A replication using both types of relations was carried out to examine these possibilities.

Link Recall. When subjects correctly verify a two-link transitive inference based on memory for recently presented set inclusion information, later recall of those relations

mediating the inference is enhanced (Sefkow, 1976). Does a similar backward review effect operate when the relations specify a linear ordering? The first condition of the present experiment tested for such an effect.

This finding rests on the assumption that both types of relations undergo similar schematization at input. While the evidence for this constructive process is strong for linear orderings, it is not so for artificial set inclusions. Nonetheless, if Potts' hypothesis is correct, then the backward review effect obtained by Sefkow using set inclusions should hold for linear orderings, at least when only correct responses are considered. If this occurs, then facilitation of the component links would have to be accounted for by any model developed to explain how such comparative decisions are made. Moreover, such a finding would support a review, as opposed to integrative, explanation of the backward review effect.

Alternatively, should component links be facilitated in the case of set inclusions but not linear orderings, Potts' hypothesis, as currently delineated, will be open to question. These results would suggest that the characterization of set inclusion schema must be reformulated or the assumption that it occurs at all rejected. Instead, the memory representation may more closely resemble or maintain the structure of the information as it was presented, in that the relations may be stored or accessed individually.

Inference Recall. In the Sefkow (1976) study, the use of the true inference probes consistently resulted in a small (about 2.5%), but nonsignificant increase in the number of inferences later recalled. The existence of such an effect will again be ascertained for both types of relations.

The two basic storage hypotheses, integrated versus propositional representation, suggest different predictions concerning the recall of inferences. Subjects are instructed to recall everything they can about the passages but are not specifically told to include all possible inferences. An implicit goal then is to reconstruct the passages as closely as possible. Regardless of the memory format, an overall advantage for the presented relations would, therefore, be expected and was obtained by Sefkow (1976) for set inclusions. If set inclusions and linear orderings are similarly represented, similar recall patterns should occur. However, if set inclusions are stored as individual propositions while the linear orderings are integrated, an interaction may well occur. That is, the percentage of adjacent relations recalled may be greater for set inclusions than linear orderings but the latter may manifeat better inferential recall.

Recognition. Subsequent to the free recall task recognition data were gathered by Sefkow (1976). The probing manipulation improved recognition of the probes but

otherwise had no effect on performance. It was collected here to replicate the truth by distance interaction for set inclusions and the standard SDE for linear orderings. Thus the assumption that subjects were at least integrating the linear orderings could be substantiated.

In addition, Sefkow asked subjects to decide whether statements they judged to be true were stated in the passages or implied. As in recent studies by Lawson (1977) and Tzeng (1975), subjects could correctly make this distinction contrary to the earlier findings of Bransford and his colleagues (e.g., Bransford & Franks, 1971). This task was included in order to compare performance on the two relations.

Presentation Order

One approach to determining whether or not any schema explanation is appropriate would be to inhibit its formation and then observe any resulting changes in probe errors and recall patterns. This manipulation is the focus of the second condition which will be discussed next.

Presentation order has been shown to be important for constructing a linear ordering from pairwise relations (Foos, Smith, Sabol & Mynatt, 1976; Huttenlocher, 1968). In the first condition of the research described above, the adjacent relations were consistently presented in a backward order as was done in the Sefkow study: DE, CD, BC, AB. Both Huttenlocher and Foos et. al. ascertained that an integrated

representation can be constructed quite easily from this sequence. Increasing the difficulty of establishing the orderings by manipulating presentation order should affect performance on the probes and, likewise, influence the backward review effect if the review effect is due to accessing an integrated representation. Even if the review effect occurs in conjunction with a proposition based representation, the presentation order manipulation may affect it by making the location and/or integration of the relevant relations more difficult. These possibilities were investigated in the second condition described here. A considerably more difficult presentation order was selected for use: CD, AB, DE, BC. It is difficult presumably because succeeding pairs add other than the optimal one new item to the ordering, thus increasing the memory load (Foos et. al., 1976).

Probes. Predictions as to the effects of presentation order on the verification of true inferences are straightforward. If true probe responses are made on the basis of an already integrated memory representation, then a significant increase in the difficulty of constructing the schema should, in turn, result in an increase in probe errors.

If individual propositions are accessed in response to a probe and then integrated, presentation order should have little effect on probe accuracy. (One caveat must be

offered. Presentation order may influence the access stage as discussed above.)

Link Recall. As was mentioned earlier, verbatim memory for prose materials like those used here can be quite good. This is true at least at the level of paraphrase and when measured soon after acquisition. Whereas some researchers (e.g., Bransford & Franks, 1971) have argued that only the overall abstracted knowledge structure is stored in memory, the Lawson (1977) and Tzeng (1975) data support the notion of two separate stores for presented and integrated information.

The verification of an inference may enhance later recall of its mediating relations as a result of accessing either memory for the individual relations or an integrated representation. Consider the case where such a backward review effect relies on an integrated structure. A basic assumption of condition 1 was that the linear orderings were being integrated. The backward review effect should then be obtained in condition 1 with both types of relation. Disruption of schema formation by the manipulation of presentation order should force subjects to rely on memory for the presented relations. The recall pattern obtained in condition 1 may then be lost for both relation types. These results would support Potts' hypothesis that set inclusions and linear orderings share an integrated form of

representation. On the other hand, enhanced recall of the mediating relations may occur because subjects access memory for these individual relations. Thus the occurrence of the review effect would be predicted in both conditions for set inclusions but would be expected for linear orderings only in condition 2 where schema formation is hindered. This pattern of results would support a propositional representation formulation for set inclusions. Differences in recall are thus expected to the degree that the effect relies on accessing an integrated schema.

A subordinate issue exists which can be clarified by the proposed manipulation of presentation order. In the Sefkow study the logical and physical order of adjacent links were confounded within the passages. One hypothesis being considered is that the facilitation of the links may be due to a directed scan of a nonschematized memory store. If the relations are stored in some propositional form maintaining order within the passage, then physical contiguity may be important. Would recall of an irrelevant relation be facilitated if its position in the passage was between the two links? This question is addressed in the present experiment.

Inference Recall. If memory for set inclusions is based upon individual propositions, changing the order in which these propositions are stored should have little effect on the recall of inferences unless order influences

access of the relevant relations. Disrupting the integration of the linear orderings should, however, cause a drop in the recall of inferences and a possible increase in adjacent relation recall. Set inclusions would likewise be affected if the integration hypothesis is correct.

Recognition. The recognition data should directly reflect the success of this manipulation. The SDE should either be greatly diminished or not obtained at all for linear orderings and minimally, inference recognition should be hurt for set inclusions too.

## Increasing the Span of the Transitive Inference Probes

The initial Sefkow (1976) study and the conditions described here so far have all used transitive inference probes that required only two relations for solution. An interesting question is: What patterns of facilitation are obtained as the span or number of component relations needed for probe verification is increased? The third and final manipulation was to compare the results of condition 1 with those obtained using more remote inferential probes.

Probes. For set inclusions, both the memory deficit and illogical processing hypotheses predict that as the span of the probes increases, so will probe error rates. However, differential results are predicted for linear orderings. As for set inclusions, a memory deficit would imply an increase

in errors while the occurrence of the SDE would predict a drop in error rates.

Link Recall. Increasing the span of the probes does not readily suggest any predictions that might help sort out the integrated versus propositional memory store controversy. Instead, it is hoped that it will help to better characterize the nature and extent of the backward review effect. Attention will be paid to the recall of component versus irrelevant false-probed relations, particularly to those component relations not sharing an end term with the probe. By involving more links in the verification of the true probes, link recall may be better than in condition 1.

Inference Recall. It will also be ascertained whether there is a tendency for more remote probes to encourage better inferential recall than in condition1. A review or strengthening account of the backward review effect would not predict this result.

Recognition. Since the verification task is basically unaffected by the probing manipulation, the results here should duplicate those of condition 1 as the identical passages were employed.

## Method

Subjects. Two hundred forty undergraduate students enrolled in introductory psychology courses at the University of Massachusetts, Amherst, served as subjects. Each received course credit for participation. The Ss were randomly assigned to small groups for testing. Within each group, Ss were randomly assigned to probe conditions with one constraint: all members of a single group listened to the same set of tape-recorded passages.

Materials. Five fictional passages were devised such that each could be presented as a five-term linear ordering or as a five-term set inclusion solely by changing the relation between four term pairings. The passages were all approximately 100 words long and dealt with a variety of topics: gardening, a primitive tribe, endangered species of fish, library books, and South American fruit growing practises. The basic structure of each passage within a condition was identical. For conditions 1 and 3 the pairwise relations were presented as in the Sefkow (1976) study: DE, CD, BC, AB. For condition 2 a more difficult order was imposed: CD, AB, DE, BC (Foos et. al., 1976; Huttenlocher, 1968). In the case of all linear orderings, unmarked comparatives were used while for set inclusions universal qualifiers were employed where A was always the smallest set and E the largest. To make the paragraphs appear more natural,

extraneous filler material was inserted between the critical statements. The four versions of each of the five topics are presented in Appendix A.

Subjects listened to all five passages, each paired with a different probe. Three true and two false probes were employed. The probes were all statements which had to be verified. Subjects were instructed to respond true, false, or ?; the latter category being reserved for the case when a subject had to guess.

The three true probes all required the verification of a transitive inference. In condition 1, the true probes were AC, BD, and CE, each composed of two relations presented adjacently in the passages. In condition 2, the probes were the same but the two component relations were not physically adjacent within the passages. Finally, in the third condition two three-link inferences (AD, BE) and one four-link inference (AE) were used. Again, the component links were presented adjacently. In all three conditions, the two false probes used to establish baseline data were irrelevant to the passages. Table 2 schematizes the passage-probe sequence for each condition.

Items for the recognition test were the 20 adjacent relations explicitly stated in the passages, the 30 true deducible relations, and 50 false items formed by reversing the two terms in each true pairing.

Table 2

Schematization of the Passages and Probes[a] as Experienced by Ss in each Condition.

| Condition 1: $O_1P_1$[b] | Condition 2: $O_2P_1$ | Condition 3: $O_1P_2$ |
|---|---|---|
| Passage 1: DE CD BC AB<br>Probe: BD? | Passage 1: CD AB DE BC<br>Probe: BD? | Passage 1: DE CD BC AB<br>Probe: BE? |
| Passage 2: DE CD BC AB<br>Probe: irrelevant false statement | Passage 2: CD AB DE BC<br>Probe: irrelevant false statement | Passage 2: DE CD BC AB<br>Probe: irrelevant false statement |
| Passage 3: DE CD BC AB<br>Probe: AC? | Passage 3: CD AB DE BC<br>Probe: AC? | Passage 3: DE CD BC AB<br>Probe: AD? |
| Passage 4: DE CD BC AB<br>Probe: irrelevant false statement | Passage 4: CD AB DE BC<br>Probe: irrelevant false statement | Passage 4: DE CD BC AB<br>Probe: irrelevant false statement |
| Passage 5: DE CD BC AB<br>Probe: CE? | Passage 5: CD AB DE BC<br>Probe: CE? | Passage 5: DE CD BC AB<br>Probe: AE? |

[a] one of the five possible probe-passage combinations---passages were always in the same order but within each condition the five probes were counterbalanced

[b] $O_1$=Adjacent Order $O_2$=Nonadjacent Order
$P_1$=Two Link Probes $P_2$=Three and Four Link Probes

Design. Forty Ss were randomly assigned to set inclusions (SI) and forty to linear orderings (LO) within each of the three conditions: (1) adjacent presentation order, two-link probes ($O_1P_1$); (2) nonadjacent presentation order, two-link probes ($O_2P_1$); and (3) adjacent presentation order, three- and four-link probes ($O_1P_2$). Passage order effects in the Sefkow (1976) study were virtually nonexistent so all Ss listened to the five passages in the same order. However, within a condition the probes were ordered according to a 5x5 Latin square (the same square was used in conditions 1 and 2, and a second square for condition 3). Each square was replicated eight times.

Procedure. All Ss were told to listen carefully to the five passages, each of which would be followed by a question testing what they had learned. Furthermore, they were informed that additional testing would follow the fifth passage (all of the instructions are presented in Appendix B).

The passages and probes were presented orally by a female, tape-recorded voice at a normal rate of approximately 135 words per minute. A click immediately followed each of the five passages. Ss were then given a sufficient time of 15 seconds to read and respond to the appropriate probe provided in a booklet. A second click terminated the response period and the next passage began. Free recall instructions followed response to the fifth passage. Each S

was given a booklet consisting of five blank sheets of paper, each headed by a key word indicating the appropriate passage to recall. The order of recall was the same as in acquisition. The Ss were instructed to "write down everything you have learned from the passages in the order indicated." They were given an adequate period of 3½ minutes per passage for recall and were told not to refer back to a recalled passage once its allotted time was up.

After the free recall test, the Ss were told that a recognition task followed. They worked through a booklet containing the recognition items at their own pace following these instructions:

> Your task is to first decide whether each sentence is true or false based on the paragraphs you heard earlier. Second, you are to rate how confident you are of that answer on a scale from one to five, where one means very low confidence and five means very high confidence. Third, if you decided the sentence was true, you must decide if it was explicitly stated or merely implied and again rate your confidence in this answer from one to five.

A 25 page response booklet was provided with four response blocks per page. No two items from the same passage occurred on the same page. The Ss were required to circle the appropriate responses in each block and were encouraged to use the full range of confidence ratings.

## Results

The important aspects of the results will be made evident by first examining the results of condition 1 alone. Condition 2 performance will then be compared with condition 1 and, in turn, condition 3 with 1. In each case attention will be paid to probe responses, particularly the comparison of LO and SI true probe error rates. Second, the existence and locus of any facilitative effects will be ascertained by comparing the recall of true-probed passage links and inferences with their false-probed passage counterparts. Finally, the recognition data will be inspected primarily to determine whether the assumptions concerning integration of the orderings in each condition can be substantiated.

### Condition 1

#### Probes

The five probes were scored as either correct or incorrect for each S. All question marks were included in the latter category. As can be seen in Table 3, the overall probe error rate was relatively high: 22.3%. Significantly more errors were made on SI probes than LO probes ($\bar{D}$=6.5%, $F(1,70)=4.75$, $p < .05$). The mean true probe error rate was 24.6% greater than that for false probes ($F(1,70)=77.79$, $p < .001$). Most important to note is the interaction of

Table 3

Condition 1: Mean Percentage of Probe Errors as a Function of Relation and Probe Types

| Probe Type | Relation | | |
|---|---|---|---|
| | LO - $O_1P_1$ | SI - $O_1P_1$ | $\bar{X}$ |
| True | 26.7 | 37.5 | 32.1 |
| False | 7.5 | 7.5 | 7.5 |
| $\bar{X}$ | 19.0 | 25.5 | 22.3 |

probe and relation types ($F(1,70)=18.06$, $p < .001$). Whereas false probe error rates were the same for the two relations, an average of 10.8% more true probe errors were made for SI than LO.

## Passage Recall

Scoring. An idea unit approach was taken for scoring the recall protocols. One unit represented each of the four basic links or relations actually present in each passage. Six additional units were assigned to the inferences. Appropriate synonyms and paraphrases were accepted along with exact replications of words and phrases. Note that no credit was given for the recall of any of the five classes unless it was mentioned correctly in the context of a link or inference. The scoring was done by one judge blind to the Ss' experimental conditions. Ten protocols were randomly selected from each of the three conditions for scoring by a second judge. Reliability of the scoring procedure was high. Agreement as to the presence or absence of the 50 idea units ranged from 92% to 100%, the mean being 97.5%.

Overall Facilitation. It was important first to determine whether verifying inferences based on passage information does generally facilitate recall of the passages for both LO and SI. Table 4 presents the relevant data.

Table 4

Condition 1: Mean Percentage of Total Links and Inferences Recalled Per S

| Relation | Probe Type | | |
|---|---|---|---|
| | True Probes | False Probes | $\bar{X}$ |
| SI - $O_1P_1$ | 28.2 | 19.4 | 24.7 |
| LO - $O_1P_1$ | 24.5 | 16.4 | 21.3 |
| $\bar{X}$ | 26.4 | 17.9 | 23.0 |

Collapsing over the two types of relations, the percentage of total links and inferences recalled from the passages was greater following true probes than false ($\bar{D}$=8.5%, $F(1,70)=27.66$, $p<.001$). This also held true for the individual relations as there was no interaction ($F(1,70)=.03$). In the past, performance on LO has consistently exceeded SI when the criterion measure was recognition (e.g., Griggs, 1976; Carroll & Kammann, 1977; etc.). Using free recall as a criterion, no significant difference was detected. In fact, SI Ss recalled 3.4% more total links and inferences than LO Ss. The probe order main effect was not significant here or in any of the subsequent analyses so no further mention will be made of it.

Link Recall: Locus of the Effect. Each S's recall of the presented relations was parsed into three categories. First, the six relations necessary for true probe verification were identified as component relations or links: AB and BC from the AC probed passage, BC and CD from the BD probed passage, and CD and DE from the CE probed passage. The remaining six relations from true probed passages, those unrelated to the verification task, were designated as irrelevant links. The final category included links recalled from the false-probed passages.

As in the Sefkow (1976) study, an adjustment of the scores was necessary prior to the analysis to account for

preexperimental differences in recall existing between the three link categories. Data from both SI and LO false-probed passages indicated that, in general, the outer links of the chains (AB, DE) were less likely to be recalled than the center links (BC, CD). The link categories were comprised of differing quantities of these two classes, resulting in the expectation that the percentage of links recalled per category would differ even prior to any experimental manipulation. To correct for these differences; i.e., equate the preexperimental expected recall levels of the three link categories, it was sufficient to compute category scores by taking the unweighted mean percentage of center and outer links recalled in each category.

The adjusted mean percentage of links recalled from the three categories can be seen in Table 5. Recall of links which had been presented in the passages was 15.7% greater for SI than LO ($F(1,70)=12.27$, $p < .001$). Collapsing over relation type, evidence for the facilitation of component link recall was obtained. The main effect due to link category was significant ($F(2,140)=10.49$, $p < .001$) as was the component versus irrelevant link contrast ($\bar{D}=9.0\%$, $t_{70}=2.76$, $p < .005$, the experimentwise error rate was held at .05). Although slightly more irrelevant links were recalled than those from false-probed passages, this difference was not significant ($t_{70}=1.09$). The interaction of relation

Table 5

Condition 1: Mean Percentage (Adjusted)

of Links Recalled Per S

| Relation | Link Type | | | |
|---|---|---|---|---|
| | Component Links | Irrelevant Links | False-Probed Passage Links | $\bar{X}$ |
| SI - $O_1P_1$ | 58.8 | 42.8 | 39.4 | 47.0 |
| LO - $O_1P_1$ | 34.7 | 32.8 | 26.3 | 31.3 |
| $\bar{X}$ | 46.8 | 37.8 | 32.9 | 39.2 |

type with the link categories is of primary interest ($F(2,140)=3.00$, $p<.05$). Component link recall was 16.0% greater than irrelevant link recall for SI ($t_{35}=3.37$, $p<.001$), but they were statistically the same for LO ($\bar{D}=1.9\%$, $t_{35}=1.47$). In fact, recall of links from LO true-probed passages was not significantly greater than from false-probed passages ($\bar{D}=6.9\%$, $t_{35}=1.82$).

The above analyses included recall data from passages whose true probes were responded to incorrectly. The source of these errors is unknown but clearly it can not be assumed that any facilitating review of memory took place or that memory representations were even available for review. The data associated with true probe errors were therefore eliminated. The revised data are in Table 6.

Enhanced recall of SI component over irrelevant links rose to 26.1%. LO component link recall also increased but the 5.4% difference between it and irrelevant link recall was not significant ($t_{39}=1.06$, $p>.20$); this increase was large enough though to make the difference between recall of LO true- and false-probed passage links meaningful ($\bar{D}=9.1\%$, $t_{39}=2.08$, $p<.025$).

In summary, true probes enhanced memory for both SI and LO relations presented explicitly in the passages. This effect was much stronger for SI than LO and was concentrated on those relations which mediated the probe inferences only in the case of SI.

Table 6

Condition 1: Mean Percentage of Links Recalled Per S Eliminating Recall Associated with True Probe Errors

| Relation | Link Type | | | |
|---|---|---|---|---|
| | Component Links | Irrelevant Links | False-Probed Passage Links | $\bar{X}$ |
| SI - $O_1P_1$ | 66.7 | 40.6 | 39.4 | 48.0 |
| LO - $O_1P_1$ | 38.1 | 32.7 | 26.3 | 31.8 |
| $\bar{X}$ | 52.4 | 36.7 | 32.9 | 39.9 |

Inference Recall. The data are presented in Table 7. Collapsing over probe type, more inferences were recalled by LO Ss than SI, though 4.4% difference was not significant ($F(1,70)=2.74$, $p > .10$). There was a main effect due to probe type ($F(1,70)=23.43$, $p < .001$). Verifying true inferences resulted in an 8.3% increase in inference recall. The true probe advantage held for both relations as the interaction did not reach significance.

Included in the above data was recall of the three true probes. By eliminating them from consideration, it can be determined whether the probing manipulation encouraged Ss to draw additional inferences. Also data associated with true probes which were responded to incorrectly can be disregarded for a clearer picture of the effects. The adjusted means are presented parenthetically in Table 7. More inferences, independent of the true probes, were recalled from LO passages than SI ($\bar{D}=7.6\%$, $t_{70}=2.50$, $p < .01$). True-probed passage recall was better than false-probed passage recall for both relations; however, this difference was significant only in the case of LO ($\bar{D}=9.5\%$, $t_{39}=2.43$, $p < .01$; SI: $\bar{D}=3.1\%$, $t_{39}=1.51$, $p < .05$). Apparently, the true probes were most effective in enhancing LO inference recall.

A final comparison can be made between recall of presented versus inferred relations. It can be seen in Figure 1 that Ss recalled substantially more presented

Table 7

Condition 1: Mean Percentage of Inferences Recalled Per S

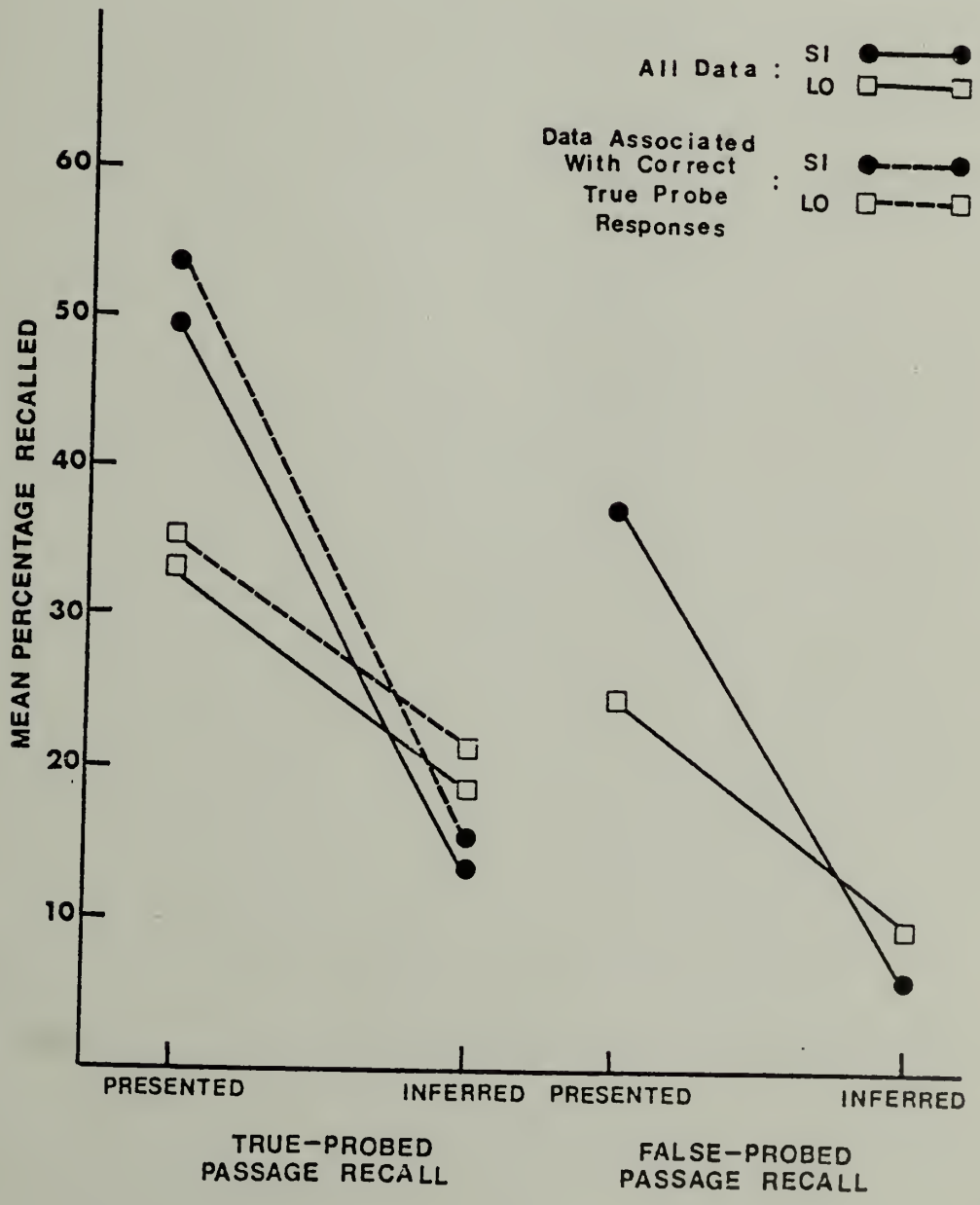
| Relation | Probe Type | | |
|---|---|---|---|
| | True Probes | False Probes | $\bar{X}$ |
| SI - $O_1P_1$ | 13.6[a] (9.1[b]) | 6.0 | 10.7 (7.9) |
| LO - $O_1P_1$ | 18.6 (19.3) | 9.8 | 15.1 (15.5) |
| $\bar{X}$ | 16.2 (14.2) | 7.9 | 12.9 (11.7) |

[a]recall of all inferences from true-probed passages

[b]disregarding true probes recalled and recall from passages whose probes were responded to incorrectly

Figure 1

Condition 1: Mean Percentage of Presented and Inferred Relations Recalled from True- and False-Probed Passages



information than inferred from both true-probed passages ($F(1,70)=115.70$, $p < .001$) and false-probed passages ($F(1,70)=102.91$, $p < .001$). Whereas the main effects due to relation type were nonsignificant, interactions were obtained (true probes: $F(1,70)=20.88$, $p < .001$; false probes: $F(1,70)=11.74$, $p < .001$). Recall of presented relations was greater for SI than LO and just the opposite occurred with inferences (only the former contrasts reached significance: true probes: $t_{78}=3.31$, $p < .01$; false probes: $t_{78}=2.99$, $p < .01$). By disregarding data associated with true probe errors, these effects were enhanced (see Figure 1).

Recognition

Scoring. Recognition data were obtained from fifteen LO and fifteen SI Ss, three from each of the five probe-order conditions. The data were collected to reflect the degree of confidence Ss had in their judgments of whether a relation was true or false based on passage information. If judged true, they further decided whether the relation had been directly stated or was logically implied. These data were scored as follows: Ss' ratings were converted into numerical values. Responses of false with confidence ratings of five through one were converted to zero through four, respectively. True responses with confidence ratings of one through five were changed to five through nine, respectively. This resulted in a ten point scale ranging from a strong

false (zero) to a strong true judgment (nine). The same scale was created for the stated-implied responses with zero representing a very confident "implied" and nine a very confident "stated" judgment.

Recognition of True- versus False-Probed Passages. Verifying true inferential probes had little discernable effect on later recognition of passage information. A series of analyses indicated that the true probes did not enhance the recognition of either links or inferences over those same items subject to false probes. The recall data demonstrated that the strength of the facilitative effect was concentrated on SI component link recall. True-false recognition ratings for these relations were actually lower, though not significantly so, than those for irrelevant or false-probed passage links ($\bar{X}$=6.37, 6.74 and 6.47, respectively). These findings replicated those of Sefkow (1976).

The true probes presented as recognition items were more confidently rated as being true ($\bar{X}_{SI}$=6.33, $\bar{X}_{LO}$=7.20) than other true, two-link nonprobe inferences ($\bar{X}_{SI}$=5.86, $\bar{X}_{LO}$=6.24; SI: $t_{28}$=2.58, $p < .01$; LO: $t_{28}$=3.71, $p < .005$). Again, similar findings were obtained by Sefkow.

The above analyses were performed on the data from conditions 2 and 3 with the same results: no differential ratings of true- versus false-probed passage relations with the exception of the true probes. No further mention of

these analyses will therefore be made.

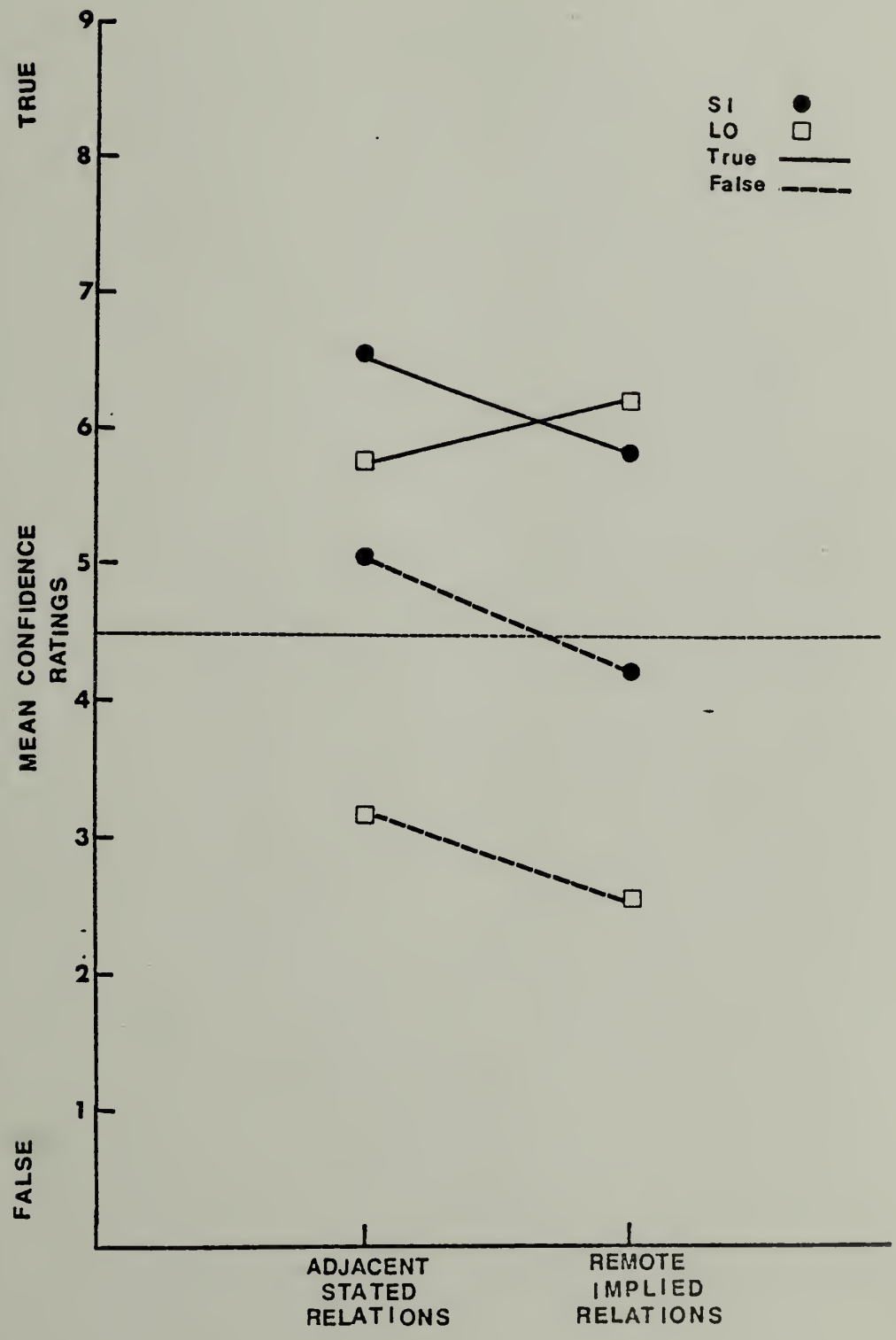
Recognition of True and False Items as a Function of Relation-type and Distance. Recall that hypotheses concerning differences in SI and LO processing are based on the finding that LO recognition accuracy for both true and false items was better on remote relations (inferences requiring two or more mediating relations) than on adjacent relations (those relations sufficient to describe the ordering). The same was found for SI false relations but just the opposite occurred on true items: adjacent relations were recognized more accurately than remote.

The true-false ratings obtained in the present study were examined for the above effects. The data for condition 1 are presented in Figure 2 and do represent a replication of the previous findings. A set of 12 contrasts were performed holding the experimentwise error rate at $\alpha = .05$ where the critical $t_{14}=2.95$ and $t_{28}=2.74$. LO true, adjacent items were rated significantly within the true range ($\bar{X} > 4.5$, $t_{14}=4.31$) and confidence in their truth increased for remotes ($t_{14}=3.08$). Likewise, LO false, adjacent relations were rated significantly within the false range ($\bar{X} < 4.5$, $t_{14}=5.28$) and confidence increased (i.e., dropped toward zero) for false remotes ($t_{14}=3.35$). Apparently, it can be assumed that Ss were integrating the LO.

True SI adjacent relations were also well within the

Figure 2

Condition 1: Mean True-False Confidence Ratings

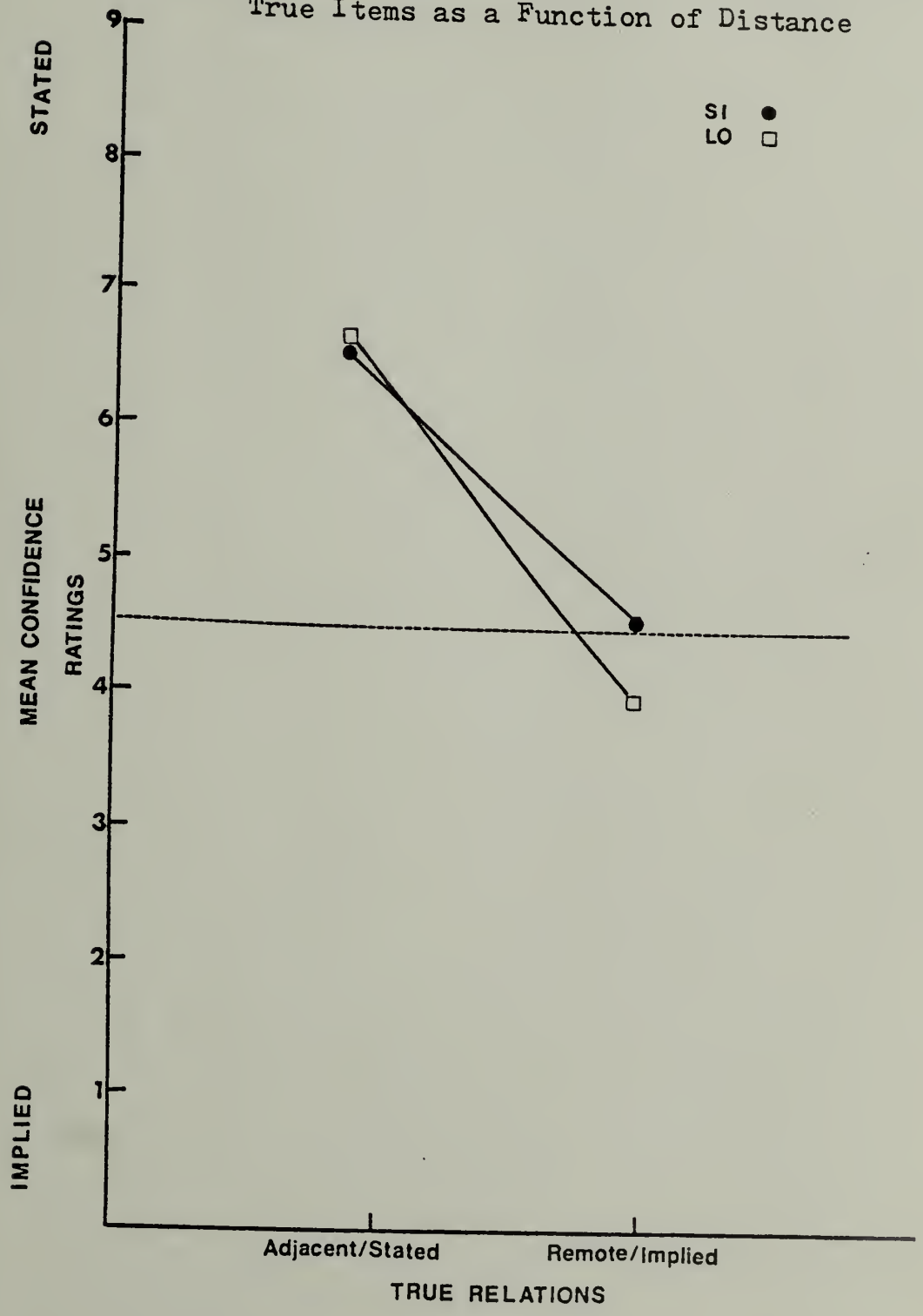


true range ($\bar{X} > 4.5$, $t_{14}=7.92$) as were the true remote relations ($t_{14}=4.89$), but confidence fell significantly for the latter ($t_{14}=3.09$). Performance on SI false items was generally poor. As in the earlier studies, Ss tended to rate false adjacent relations as true (though not significantly above 4.5, $t_{14}=1.73$). The remote ratings did decrease ($t_{14}=4.18$) but still were not firmly within the false range ($t_{14} < 1$). Note that the recognition of true items paralleled the recall data. Whereas accuracy was better for SI than LO on adjacent true items, the opposite was true for remotes (neither contrast was significant however: $t_{28}=2.14$ and $t_{28}=1.16$, respectively).

Stated-Implied Judgments. Mean confidence ratings for the stated-implied judgments for true relations are plotted in Figure 3. An analysis of variance revealed only a main effect due to distance ($F(1,28)=16.73$, $p < .001$). Given that the relations were judged true, Ss were equally accurate in identifying true LO and SI adjacent relations as "stated": both means were significantly within the "stated" range ($\bar{X} > 4.5$: LO, $t_{14}=7.63$, $p < .001$; SI, $t_{14}=7.24$, $p < .001$). Confidence ratings for the remote relations dropped toward the "implied" range, more so for LO than SI though this interaction was not significant. Ss could then distinguish inferences from stated facts but were not so confident in their categorization of the former. Only the LO mean rating

Figure 3

Condition 1: Mean Stated-Implied Confidence Ratings for True Items as a Function of Distance



was within the implied range and then not significantly so.

## Condition 2

### Probes

The mean true probe error rates are shown in Table 8. Overall error rates did not differ as a function of presentation order or type of relation though their interaction was highly significant ($F(1,156)=10.86$, $p < .001$). When the presentation order allowed relatively easy integration of the relations (condition 1), 10.8% fewer errors were made in verifying LO inferential probes than SI. As integration was made more difficult (condition 2), there was a 14.1% increase in LO errors while the error rate actually dropped 8.3% for SI. In contrast to the easier presentation order, LO errors now exceeded SI errors by 11.6%. Except for the drop in SI errors from condition 1 to 2, all of these comparisons were substantiated by the Newman-Keuls procedure ($S_{\bar{y}} = .10$, $\alpha = .05$).

False probe errors were marginal in condition 2. SI Ss made no errors while only 1.3% were made by those listening to LO passages.

### Passage Recall

Scoring. All scoring was done as in condition 1.

Overall Facilitation. Table 9 presents the relevant

Table 8

Condition 1 versus 2: Mean Percentage of True Probe Errors

| Condition | Relation | | |
|---|---|---|---|
| | SI | LO | $\bar{X}$ |
| Condition 1 - $O_1P_1$ | 37.5 | 26.7 | 32.1 |
| Condition 2 - $O_2P_1$ | 29.2 | 40.8 | 35.0 |
| $\bar{X}$ | 33.4 | 33.8 | 33.6 |

Table 9

Condition 1 versus 2: Mean Percentage of Total Links and Inferences Recalled Per S

| Condition/ Relation | Probe Type | | |
|---|---|---|---|
| | True Probes | False Probes | $\bar{X}$ |
| | Adjacent Presentation Order | | |
| SI - $O_1P_1$ | 28.2 | 19.4 | 24.7 |
| LO - $O_1P_1$ | 24.5 | 16.4 | 21.3 |
| $\bar{x}$ | 26.4 | 17.9 | 23.0 |
| | Nonadjacent Presentation Order | | |
| SI - $O_2P_1$ | 26.3 | 21.4 | 24.3 |
| LO - $O_2P_1$ | 11.1 | 8.8 | 10.2 |
| $\bar{x}$ | 18.7 | 15.1 | 17.3 |
| $\bar{X}$ | 22.5 | 16.5 | 20.1 |

data. Though recall in condition 2 was enhanced by the true probes ($t_{70}=3.48$, $p<.001$), it was to a lesser extent than in condition 1: a significant probe type by condition interaction was obtained ($F(1,140)=6.38$, $p<.025$). The condition 2 true probe recall advantage appears to be stronger for SI than LO. This interaction was not significant. Overall recall was lower in condition 2 than 1 ($F(1,140)=11.62$, $p<.001$) due to an 11.1% drop in LO performance. SI recall remained virtually the same, thus the significant condition by relation interaction ($F(1,140)=10.43$, $p<.005$). Likewise, the detrimental effect that the change in presentation order had on LO accounted for the overall advantage of SI recall ($F(1,140)=27.69$, $p<.001$). To review, SI Ss recalled approximately the same amount of information given both presentation orders with the difference between true- and false-probed passages being smaller for the nonadjacent order. The latter was also true for LO but the nonadjacent presentation order caused a substantial overall drop in LO recall.

Link Recall: Locus of the Effect. As in condition 1, recall of the relations actually presented in the passages was parsed into three categories: component relations necessary for true probe verification, relations from true-probed passages irrelevant to the true probes, and relations from false-probed passages. These data were also adjusted

for preexperimental differences in recall as was done in condition 1.

The data from conditions 1 and 2 are presented in Table 10 for comparison. An overall analysis of variance indicated that the recall advantage of SI over LO evident in condition 1 was also significant in condition 2 (main effect due to relation type: $F(1,140)=49.16$, $p < .001$: no relation by presentation order interaction: $F(1,140)=2.42$, $p > .10$). Recall of presented information was better given the adjacent presentation order than the nonadjacent order ($\bar{D}=7.9\%$, $F(1,140)=7.35$, $p < .01$). By increasing the difficulty of integrating LO adjacent relations one would expect a drop in inference recall. The effect was more pervasive than that: the recall level for presented relations fell. Whereas the drop in recall from condition 1 to 2 was 3.2% for SI, it was nearly four times greater for LO ($\bar{D}=12.4\%$). Integration thus seems to be a necessary prerequisite for successful recall of LO adjacent relations. The link category main effect was highly significant ($F(2,280)=24.11$, $p < .001$). As reported earlier, component link recall was enhanced only for SI in condition 1. Inspection of the data for condition 2 shows this facilitation for both LO and SI. The SI component versus irrelevant link contrast was significant ($\bar{D}=15.3\%$, $t_{35}=3.07$, $p < .001$: experimentwise error rate for the set of contrasts was held at .05). The

Table 10

Condition 1 versus 2: Mean Percentage (Adjusted) of Links Recalled Per S

| Condition/ Relation | Link Type | | | |
|---|---|---|---|---|
| | Component Links | Irrelevant Links | False-Probed Passage Links | $\bar{X}$ |
| | Adjacent Presentation Order | | | |
| SI - $O_1P_1$ | 58.8 | 42.8 | 39.4 | 47.0 |
| LO - $O_1P_1$ | 34.7 | 32.8 | 26.3 | 31.3 |
| $\bar{x}$ | 46.8 | 37.8 | 32.9 | 39.2 |
| | Nonadjacent Presentation Order | | | |
| SI - $O_2P_1$ | 54.4 | 39.1 | 37.5 | 43.7 |
| LO - $O_2P_1$ | 25.0 | 15.9 | 15.9 | 18.9 |
| $\bar{x}$ | 39.7 | 27.5 | 26.7 | 31.3 |
| $\bar{X}$ | 43.2 | 32.7 | 29.8 | 35.2 |

Table 11

Condition 1 versus 2: Mean Percentage of Links Recalled Per S Eliminating Recall Associated with True Probe Errors

| Condition/ Relation | Link Type | | | |
|---|---|---|---|---|
| | Component Links | Irrelevant Links | False-Probed Passage Links | $\bar{X}$ |
| | Adjacent Presentation Order | | | |
| SI - $O_1P_1$ | 66.7 | 40.6 | 39.4 | 48.0 |
| LO - $O_1P_1$ | 38.1 | 32.7 | 26.3 | 31.8 |
| $\bar{x}$ | 52.4 | 36.7 | 32.9 | 39.9 |
| | Nonadjacent Presentation Order | | | |
| SI - $O_2P_1$ | 52.7 | 35.7 | 37.5 | 41.5 |
| LO - $O_2P_1$ | 27.4 | 14.3 | 15.9 | 18.9 |
| $\bar{x}$ | 40.1 | 25.0 | 26.7 | 30.2 |
| $\bar{X}$ | 46.2 | 30.9 | 29.8 | 35.1 |

9.1% difference between LO component and irrelevant link recall fell just short of significance ($t_{35}=2.56$). Finally, irrelevant link recall was not significantly different from that of false-probed passage links.

Recall of the presented, adjacent relations was reexamined, disregarding data associated with true probe errors. The revised means for both conditions can be seen in Table 11. Enhanced recall of condition 2 component over irrelevant links rose to 17.0% for SI and a now significant 13.1% for LO ($t_{38}=2.97$, $p<.001$).

The most important findings from these data were that first, presentation order and hence ease of integration had little if any effect on the pattern of facilitated recall for SI. Second, the same pattern was obtained with LO only when initial integration was hampered via a difficult presentation order.

In condition 2, the adjacent relations were ordered in each passage such that either one or two relations were presented between the two needed to mediate the true probe inferences (see Table 2, condition 2: $O_2P_1$). Recall of these intervening items was compared to recall of the appropriate relations from false-probed passages. This was done to determine whether any facilitation occurred, possibly due to a scanning process initiated by the true probes. The data are presented in Table 12. The small differences in recall were neither consistent in direction

Table 12

Condition 2: Mean Percentage of Links Recalled within Physical Span of True Probes

| Relation | Link Categories | | |
|---|---|---|---|
| | Recall of single intervening relations | Recall of double intervening relations | Recall of corresponding false items |
| SI - $O_2P_1$ | 40.0[a] (40.8[b]) | 33.8 (37.2) | 41.9 |
| LO - $O_2P_1$ | 15.0 (14.3) | 13.8 (12.8) | 13.1 |

[a]recall of all possible relations

[b]recall of relations disregarding those associated with true-probe errors

nor significant in size. This was true for all the data. It was also true when only recall was considered from passages whost true probes were responded to correctly. If some type of ordered scan occurred, it was not responsible for enhancing recall.

Inference Recall. The data for conditions 1 and 2 are summarized in Table 13. In general, true probes resulted in better inference recall than false probes ($\bar{D}$=4.9%, $F(1,140)$= 20.79, $p < .001$). Probe type also interacted with condition ($F(1,140)=10.26$, $p < .005$): true probes facilitated the recall of inferences for both relations in condition 1 ($t_{70}=4.84$, $p < .001$) but for neither in condition 2 ($t_{70}=1.13$, $p > .10$; condition by relation and condition by relation by probe type: $F(1,140) < 1$).

The degree to which the true probes facilitated recall of other inferences was assessed by eliminating recall associated with true probe errors and recall of the true probes themselves. These adjusted means are presented parenthetically in Table 13. The true probes encouraged a significant amount of additional inference recall only for LO where the adjacent relations were easily integrated.

Consider recall of all possible inferences collapsed over probe type. More inferences were recalled from the easily integrated, adjacent presentation order than the nonadjacent order ($\bar{D}$=4.6%, $F(1,140)=16.16$, $p < .001$). This

Table 13

Condition 1 versus 2: Mean Percentage of Inferences Recalled Per S

| Condition/ Relation | Probe Type | | |
|---|---|---|---|
| | True Probes | False Probes | $\bar{X}$ |
| | Adjacent Presentation Order | | |
| SI - $O_1P_1$ | 13.8[a] (9.1[b]) | 6.0 | 10.7 (7.9) |
| LO - $O_1P_1$ | 18.6 (19.3) | 9.8 | 15.1 (15.5) |
| $\bar{x}$ | 16.2 (14.2) | 7.9 | 12.9 (11.7) |
| | Nonadjacent Presentation Order | | |
| SI - $O_2P_1$ | 13.5 (11.7) | 10.8 | 12.4 (11.3) |
| LO - $O_2P_1$ | 4.2 (3.7) | 3.9 | 4.1 (3.8) |
| $\bar{x}$ | 8.8 (7.7) | 7.4 | 8.3 (7.6) |
| $\bar{X}$ | 12.5 (11.0) | 7.6 | 10.5 (9.7) |

[a]recall of all inferences from true-probed passages

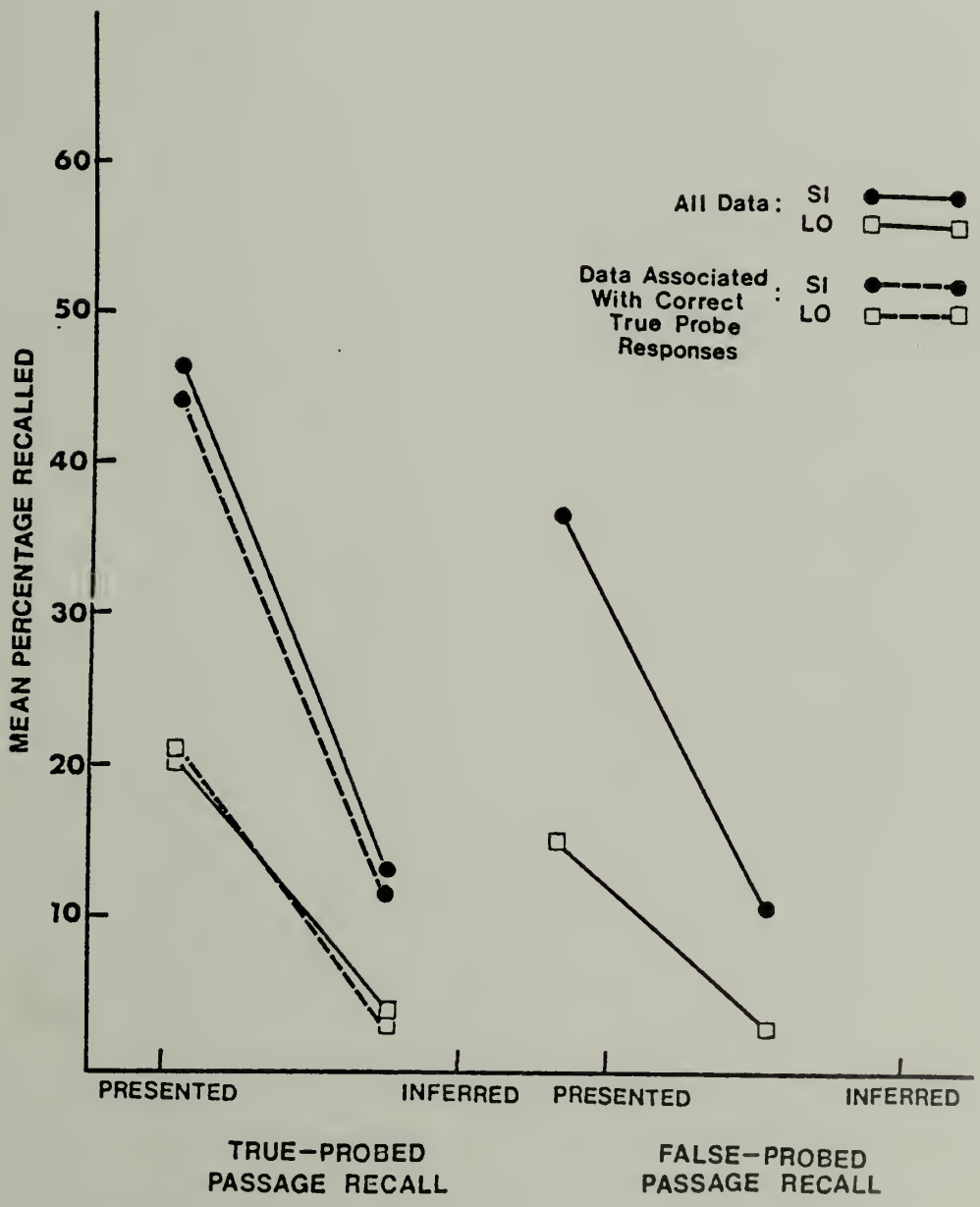
[b]recall of inferences disregarding true probes recalled and eliminating recall associated with true probe errors

occurred because of an 11.0% drop in LO recall from condition 1 to 2 ($t_{140}=6.76$, $p<.001$). SI recall actually increased slightly under the more difficult circumstances but not significantly so. As a result, SI inference recall was 8.3% greater than LO in condition 2 ($t_{140}=5.36$, $p<.001$) and a condition by relation interaction was obtained ($F(1,140)=30.70$, $p<.001$). Overall, there was no main effect due to relation type.

A last comparison to be made is between the recall of presented and inferred relations. The data for condition 2 are presented in Figure 4. As in condition 1, considerably more of the stated, adjacent relations were recalled than inferences from both true-probed passages ($F(1,70)=117.02$, $p<.001$) and false-probed passages ($F(1,70)=108.20$, $p<.001$). The main effects due to relation type were also significant (true probes: $F(1,70)=17.01$, $p<.001$; false probes: $F(1,70)=15.89$, $p<.001$). Most importantly, relation type did not interact with the adjacent-remote variable. SI recall of both presented and inferred relations exceeded LO recall of the same items. Exclusion of data associated with true-probe errors did not change the results (see Figure 4). In making the passages harder to integrate, LO not only lost its inference recall advantage but recall of the presented information dropped relative to SI also.

Figure 4

Condition 2: Mean Percentage of Presented and Inferred Relations Recalled from True- and False-Probed Passages



## Recognition

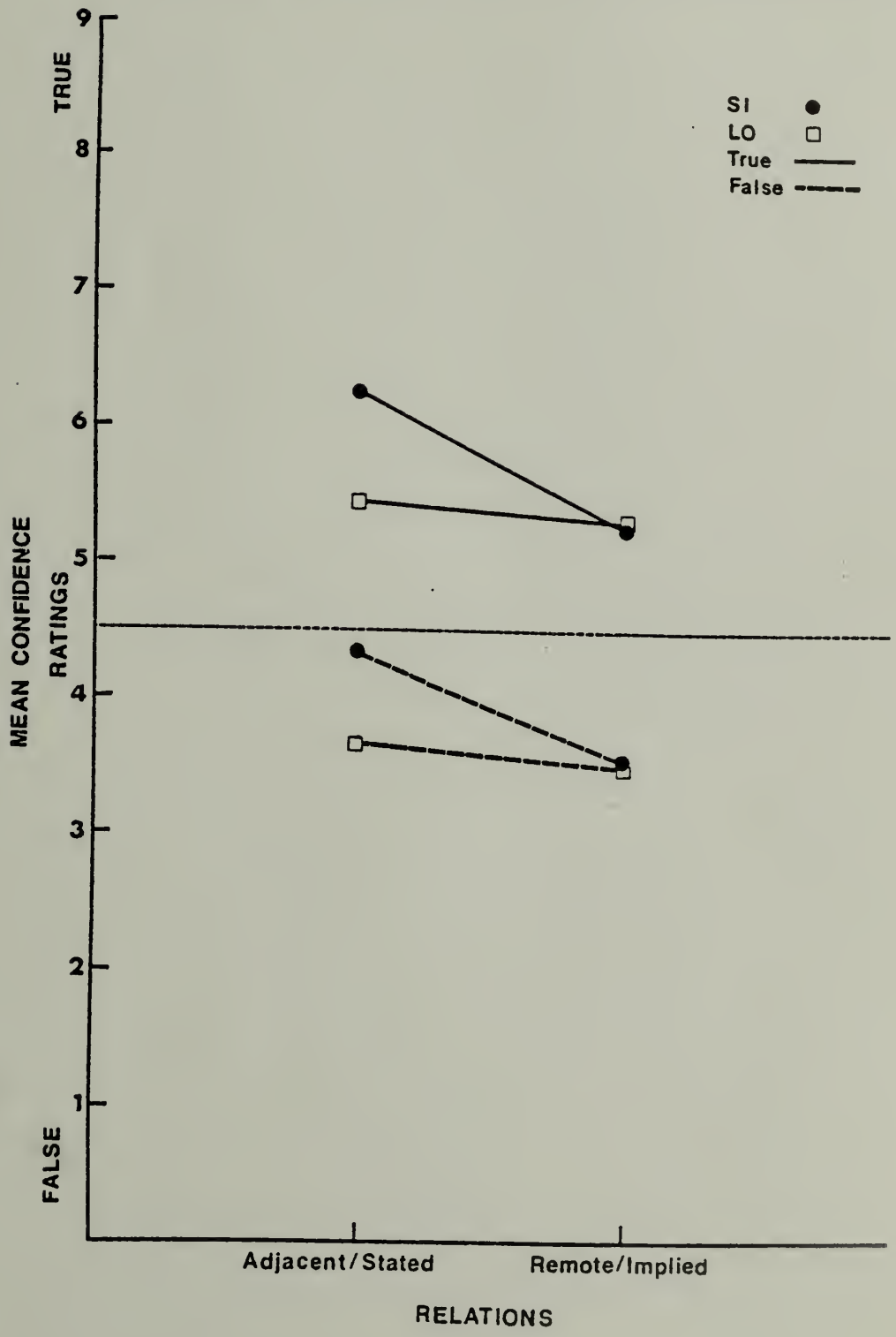
Scoring. Data were obtained from all eighty Ss in condition 2. The true-false and stated-implied judgments were converted into values on the ten point scales as described in condition 1.

Recognition of True and False Items as a Function of Relation-type and Distance. The mean true-false ratings for condition 2 are presented in Figure 5. Hampering integration by manipulating presentation order had little effect on the recognition of SI relations (cf. Figure 2). A set of 16 contrasts were performed holding the experimentwise error rate at $\alpha = .05$ where the critical $t_{39}=2.86$ and $t_{78}=2.81$. Both true SI adjacent and remote relations were rated as true ($\bar{X} > 4.5$: $t_{39}=6.25$ and $t_{39}=4.02$, respectively) with the usual drop in performance from adjacents to remotes ($t_{39}=3.24$). Accuracy on SI false items appeared better than in condition 1. The mean rating for false SI adjacent items was within the false range though not significantly so ($\bar{X} < 4.5$, $t_{39} < 1$). Ss could better identify false SI inferences ($t_{39}=4.86$), which were significantly rated as false ($\bar{X} < 4.5$, $t_{39}=4.62$).

As predicted, the difficult presentation order interfered with the integration of LO relations. The SDE was not obtained: performance was not better on remote relations than adjacent for either true or false items

Figure 5

Condition 2: Mean True-False Confidence Ratings

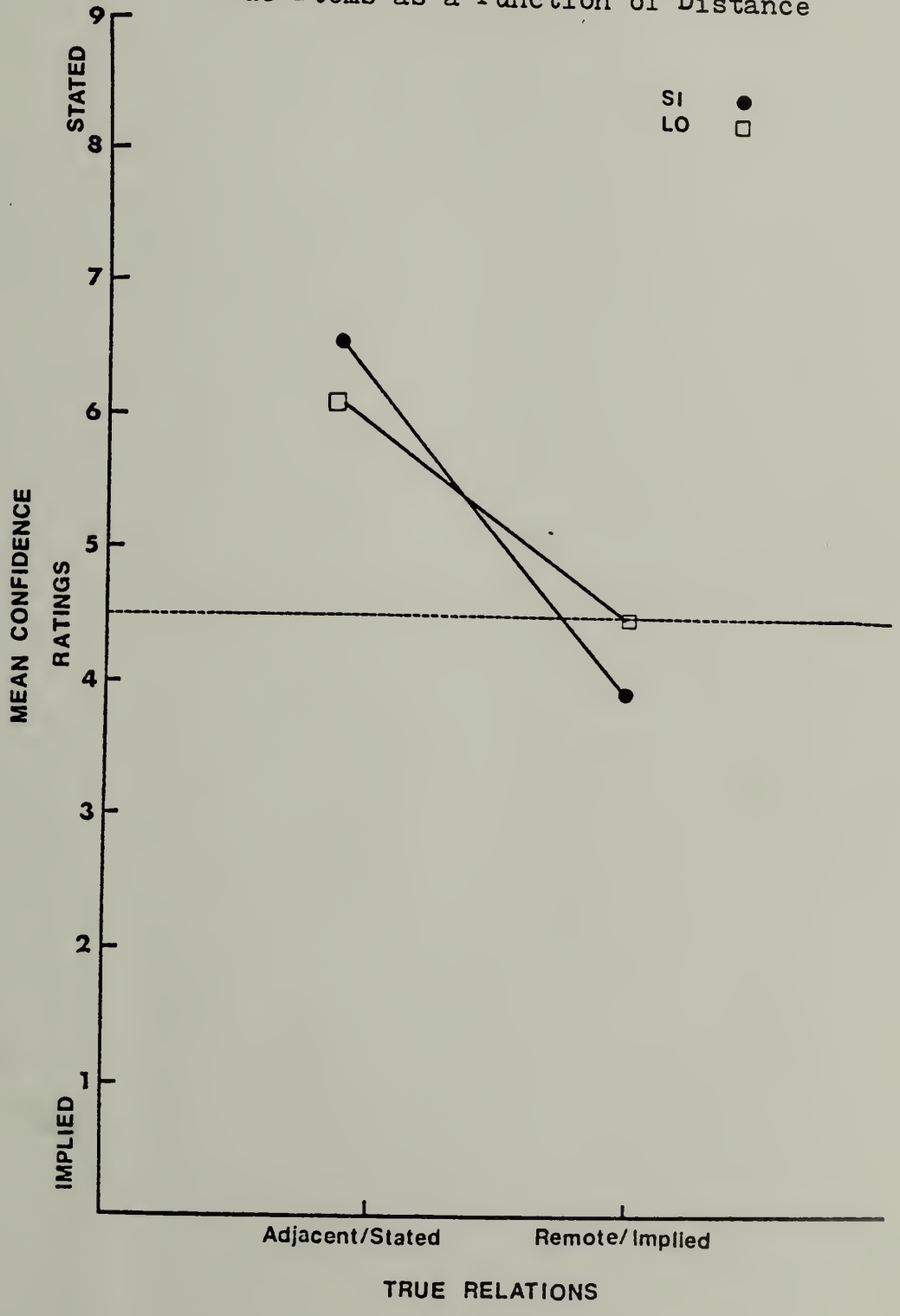


(both contrasts $t_{39} < 1$). Ss did tend to recognize all the true LO relations as true and the false relations as false ($\bar{X} > 4.5$, true adjacent: $t_{39}=3.16$, true remote: $t_{39}=3.40$; $\bar{X} < 4.5$, false adjacent: $t_{39}=2.97$, false remote: $t_{39}=2.10$). The presented SI adjacent items were better judged true than LO, while LO false adjacents were more confidently rated false ($t_{78}=1.87$ and $t_{78}=1.60$, respectively). Performance on the remote relations did not differ as a function of relation type (both contrasts $t_{78} < 1$). Ss were equally adept at recognizing the inferences.

Stated-Implied Judgments. Mean stated-implied confidence ratings for the true relations are plotted in Figure 6. SI adjacent and remote relations were more accurately recognized than the LO counterparts but this interaction was not significant. As in condition 1, the only significant effect was due to distance. Ratings for the presented, adjacent relations were higher than for remote relations ($F(1,78)= 17.82$, $p < .001$). Ss accurately rated the presented information for both relations ($\bar{X} > 4.5$: SI, $t_{39}=5.51$, $p < .001$; LO, $t_{39}=4.60$, $p < .001$). Ss were not so confident that they had inferred the remote relations. Neither mean was significantly below the guessing level of 4.5.

Figure 6

Condition 2: Mean Stated-Implied Confidence Ratings for True Items as a Function of Distance



## Condition 3

### Probes

The true probe error rates for conditions 1 and 3 are presented in Table 14. Collapsing over relations, true probe error rates did not vary significantly with condition ($F(1,156) < 1$). Substantially fewer errors were made on LO probes than SI ($\bar{D}$=19.6%, $F(1,156)=26.03$, $p < .001$). When the transitive inference probes required two mediating relations (condition 1), fewer verification errors were made for LO than SI. When the number of mediating relations increased to three and four (condition 3), the incidence of LO errors dropped while SI errors increased, thus a significant interaction was obtained ($F(1,156)=5.22$, $p < .025$). The Newman-Keuls procedure was used to show that all four error rates differed significantly ($S_{\bar{y}}=.11$, $\alpha=.05$).

As in the other conditions, few false probe errors were made. No errors occurred in conjunction with SI passages and only 2.5% with LO passages.

### Passage Recall

Scoring. All scoring was done as in condition 1.

Overall Facilitation. Table 15 shows the mean percentage of all relations, presented and inferred, recalled from true-probed and false-probed passages for conditions 1 and 3. Averaged across conditions, recall of

Table 14

Condition 1 versus 3: Mean Percentage of True Probe Errors

| Condition | Relation | | |
|---|---|---|---|
| | SI | LO | $\bar{X}$ |
| Condition 1 - $O_1P_1$ | 37.5 | 26.7 | 32.1 |
| Condition 3 - $O_1P_2$ | 44.2 | 15.8 | 30.0 |
| $\bar{X}$ | 40.9 | 21.3 | 31.1 |

Table 15

Condition 1 versus 3: Mean Percentage of Total Links and Inferences Recalled Per S

| Condition/ Relation | Probe Type | | |
|---|---|---|---|
| | True Probes | False Probes | $\bar{X}$ |
| | Two Link Probes | | |
| SI - $O_1P_1$ | 28.2 | 19.4 | 24.7 |
| LO - $O_1P_1$ | 24.5 | 16.4 | 21.3 |
| $\bar{x}$ | 26.4 | 17.9 | 23.0 |
| | Three and Four Link Probes | | |
| SI - $O_1P_2$ | 26.7 | 23.3 | 25.3 |
| LO - $O_1P_2$ | 30.3 | 30.5 | 30.4 |
| $\bar{x}$ | 28.5 | 26.9 | 27.9 |
| $\bar{X}$ | 27.4 | 22.4 | 25.5 |

true-probed passages exceeded that of false-probed passages by 5.0% ($F(1,140)=14.54$, $p < .001$). Probe type interacted with condition ($F(1,140)=6.72$, $p < .025$). Two link probes enhanced recall by 8.5% ($t_{70}=5.26$, $p < .001$) but the three and four link probes were not responsible for a significant enhancement ($\bar{D}=1.6\%$, $t_{70}=1.04$, $p > .20$). Overall, the same levels of recall were evident for SI and LO ($F(1,140) < 1$) and relation type did not interact with any other variable. Despite no enhancement due to true probes, recall was 4.9% better in condition 3 than 1 ($F(1,140)=4.48$, $p < .05$).

Inspection of the data suggests that the major difference between the two conditions was not so much in true-probed passage recall but in false-probed passage recall, particularly for LO. (The Scheffe procedure for posteriori contrasts did not substantiate any of the four comparisons, though comparison of LO false-probed passage recall for each condition fell just short of significance.) If these differences were meaningful, they could have been due to some orienting task effect (cf. Rothkopf, 1966) such that the more difficult true probes encouraged better initial processing of all ensuing passages, including those paired with false probes. If so, recall of first passages which were followed by false probes should be equivalent for conditions 1 and 3 as they represent identical circumstances. In fact, condition 3 recall was 9.0% greater than

condition 1 ($t_{62}=2.52$, $p < .02$), refuting the forward effect hypothesis.

Link Recall. Recall of the presented relations from true- and false-probed passages for conditions 1 and 3 can be compared in Table 16. Recall of SI links was 11.2% greater than LO links ($F(1,140)=11.92$, $p < .001$). The advantage for SI held up across conditions as no relation type by condition interaction was obtained ($F(1,140)=1.48$, $p > .20$). There was no significant difference in overall link recall between the two conditions ($F(1,140)=2.56$, $p > .10$). In particular, probes involving three and four mediating relations did not result in significantly better recall of links than those involving two. In general, link recall was enhanced by the true probes ($\bar{D}=4.9$, $F(1,140)=6.19$, $p < .025$), but a condition by probe type interaction was obtained ($F(1,140)=3.93$, $p < .05$). Whereas 8.8% more links were recalled from true- than false-probed passages in condition 1 ($t_{70}=3.56$, $p < .005$), only 1% more were recalled in condition 3 ($t_{70} < 1$). No other interactions reached significance.

Condition 3 data associated with true-probe errors was eliminated. The adjusted means are shown parenthetically in Table 16. Recall increased such that a significant facilitation of SI true-probed passage link recall was manifest ($\bar{D}=9.4\%$, $t_{35}=1.85$, $p < .05$). Though more links were

Table 16

Condition 1 versus 3: Mean Percentage of Links Recalled Per S

| Condition/ Relation | Probe Type | | | | $\bar{X}$ |
|---|---|---|---|---|---|
| | True Probes | | False Probes | | |
| | Two Link Probes | | | | |
| SI - $O_1P_1$ | 49.8[a] | (53.7[b]) | 39.4 | | 45.6 |
| LO - $O_1P_1$ | 33.3 | (35.4) | 26.3 | | 30.5 |
| $\bar{x}$ | 41.6 | | 32.9 | | 38.1 |
| | Three and Four Link Probes | | | | |
| SI - $O_1P_2$ | 48.3 | (52.8) | 44.7 | (43.4) | 46.9 |
| LO - $O_1P_2$ | 39.0 | (42.9) | 40.6 | (41.3) | 39.6 |
| $\bar{x}$ | 43.7 | | 42.7 | | 43.3 |
| $\bar{X}$ | 42.6 | | 37.7 | | 40.7 |

[a]recall of all links from true probed passages

[b]recall of links disregarding those associated with true-probe errors

involved in the mediation of the probe inferences, the enhancement was not greater than in condition 1. The revised condition 3 LO means show only a very slight, nonsignificant increase in link recall following true probes ($\bar{D}$=1.6%, $t_{38} < 1$). LO link recall was enhanced by the verification of two link inferences, but did not occur with the more extensive inferential probes of condition 3.

In hopes of more carefully identifying the locus of the effect, the revised link recall was further broken down into several categories: (1) component or mediating links sharing an end term (ME) with the probes (AB and CD for probe AD, BC and DE for BE, and AB and DE for AE), (2) mediating links not sharing an end term ($M\bar{E}$) with the probes (BC for probe AD, CD for probe BE, and BC and CD for probe AE), and (3) irrelevant or nonmediating true-probed passage links (I). Refer to the inset in Figure 7. Again note that recall of links from false-probed passages tends to vary such that outer links of a passage (AB,DE) are less likely to be recalled than center links (BC,CD). Whereas all four links are represented in the ME category, $M\bar{E}$ links are exclusively from the center and I links are exclusively from the outer category. Therefore, while ME link recall was compared to overall false-probed passage link recall, $M\bar{E}$ and I links were compared with the recall of center and outer false-probed passage links, respectively.

It can be seen from Figure 7 that more SI links were

Figure 7

Condition 3: Set Inclusion Link Recall From Passages Whose Probes Were Responded to Correctly

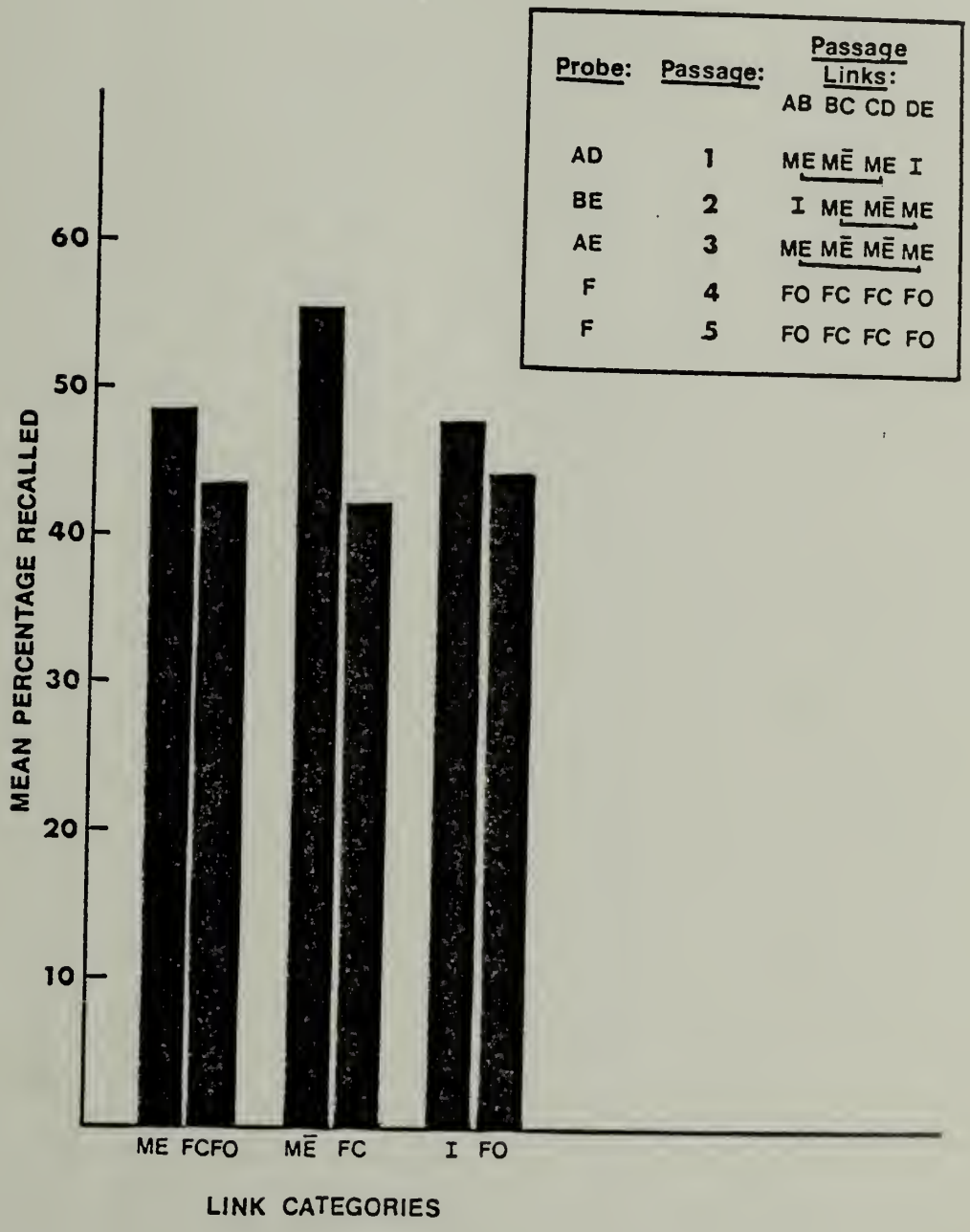
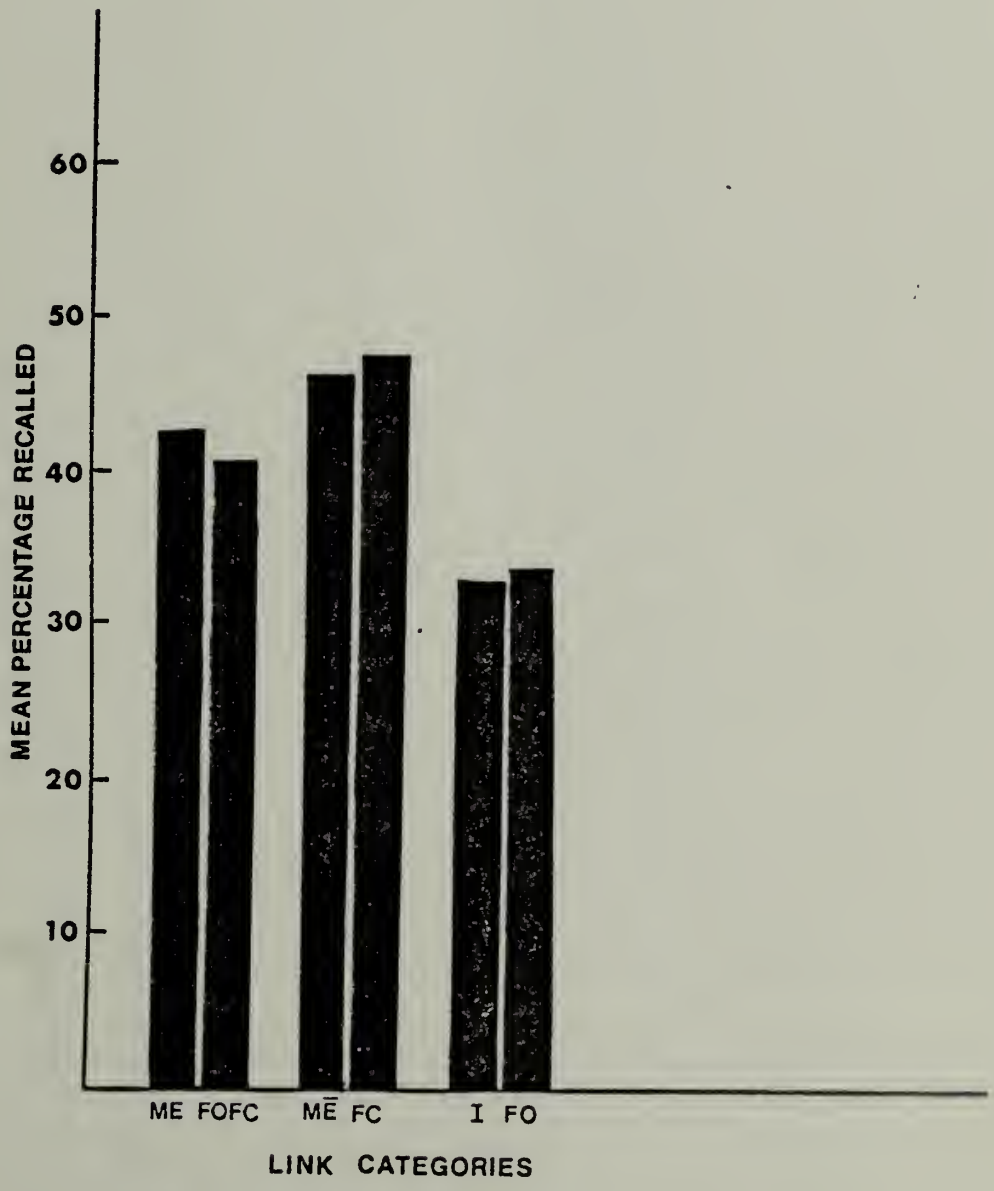


Figure 8

Condition 3: Linear Ordering Link Recall From Passages Whose Probes Were Responded to Correctly



recalled from all three true-probed SI passage categories than their false-probed counterparts; however, only the $M\bar{E}$ versus FC contrast was statistically significant ($t_{35}=2.31$, $p < .025$). LO means for the same categories are shown in Figure 8. Only in the case where links share an end term with the true probes (ME) was recall enhanced and then not significantly so.

To review, the three and four link probes of condition 3 were responsible for enhancing the recall of the innermost mediating SI links and possibly to some extent all other true-probed passage SI links. No such pattern of facilitation was obtained with LO.

Inference Recall. The mean percentage of inferences recalled in conditions 1 and 3 as a function of probe and relation type are presented in Table 17. Overall, 4.7% more inferences were recalled in condition 3 than 1 ($F(1,140)=12.01$, $p < .001$). Moreover, LO inference recall exceeded SI by 8.9% and 5.1% more inferences were recalled from true- than false-probed passages ($F(1,140)=13.89$, $p < .001$; $F(1,140)=15.38$, $p < .001$; respectively). The interactions were of particular interest. The true versus false probe advantage held only for condition 1 ($F(1,140)=5.29$, $p < .025$; condition 3: true-probed recall equalled false-probed recall, $t_{70}=1.04$, $p > .10$). Apparently the more extensive inferential probes did not effectively encourage

Table 17

Condition 1 versus 3: Mean Percentage of Inferences Recalled Per S

| Condition/ Relation | Probe Type | | | |
|---|---|---|---|---|
| | True Probes | | False Probes | $\bar{X}$ |
| | Two Link Probes | | | |
| SI - $O_1P_1$ | 13.8[a] | (9.1[b]) | 6.0 | 10.7 |
| LO - $O_1P_1$ | 18.6 | (19.3) | 9.8 | 15.1 |
| $\bar{x}$ | 16.2 | | 7.9 | 12.9 |
| | Three and Four Link Probes | | | |
| SI - $O_1P_2$ | 12.2 | (13.7) | 9.2 | 11.0 |
| LO - $O_1P_2$ | 24.6 | (27.1) | 23.7 | 24.3 |
| $\bar{x}$ | 18.4 | | 16.5 | 17.6 |
| $\bar{X}$ | 17.3 | | 12.2 | 15.3 |

[a]recall of all inferences from true-probed passages

[b]recall of inferences disregarding true probes recalled and eliminating recall associated with true probe errors

better inference recall for either LO or SI (no significant second-order interaction). Condition also interacted with relation type ($F(1,140)=3.85$, $p<.05$). While overall LO and SI recall were the same in condition 1, LO exceeded SI by 13.3% in condition 3 ($t_{70}=3.65$, $p<.005$). Contrasting the four cells of condition 1 with the four of condition 3, the only significant difference was between the LO false-probed passage cells. One possible explanation for the difference in recall levels could be that different S response criteria were in use. If so, condition 3 would also result in an increase in the recall of inaccurate and incomplete information. The mean percentage of single class elements recalled but not as part of an adjacent relation or inference, as well as incorrect relation reversals (e.g., CcB instead of BcC) are presented in Table 18. Differential response criteria do not appear to be operating. Condition 1 and 3 recall of this type was substantially the same. Once again there is no good theoretical account as to why false-probed passage recall should vary across these conditions other than an obvious S difference.

Recall of the true probes and relations associated with true probe errors were disregarded (means presented parenthetically in Table 17). There was a trend toward better inference recall following true probes but significance was obtained only for condition 1 LO recall. The major finding here was that as the number of mediating links

Table 18

Mean Percentage of Single Elements[1] and Erroneous Relation Reversals[2] Recalled per Condition

| Condition | Relation | | | |
|---|---|---|---|---|
| | SI | | LO | |
| | Single Elements | Reversals | Single Elements | Reversals |
| Condition 1: $O_1$ | 6.0 | 1.05 | 9.6 | .85 |
| Condition 2: $O_2$ | 5.6 | .75 | 19.5 | 1.30 |
| Condition 3: $O_1$ | 7.4 | .95 | 9.9 | .95 |

[1] based upon a possible 1000 single elements: 40 Ss x 5 passages x 5 elements

[2] based upon a possible 2000 reversed relations: 40 Ss x 5 passages x 10 relations

involved in the verification of the probes increased, better inference recall did not occur.

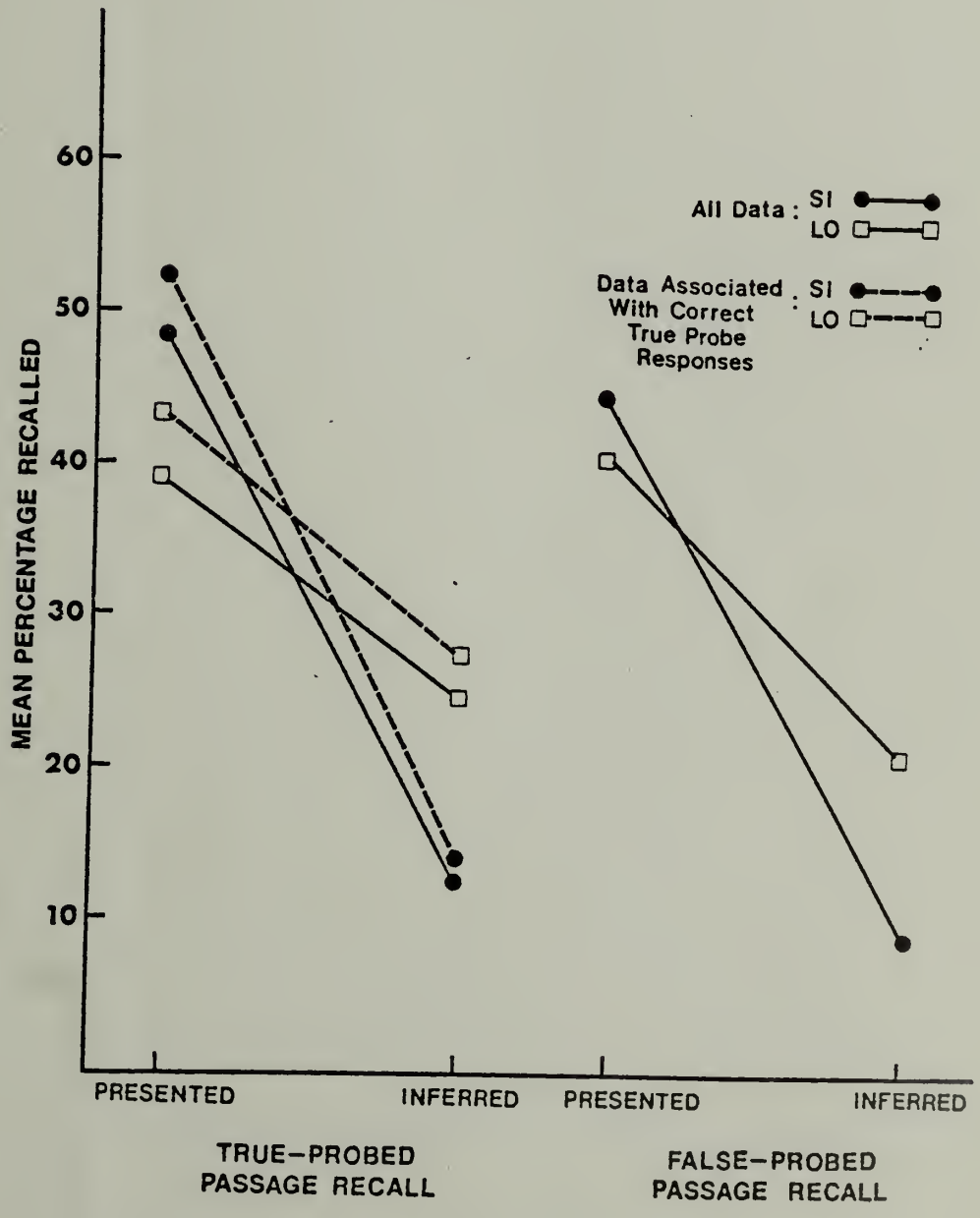
The mean recall of the presented, adjacent relations and the inferred, remote relations are plotted in Figure 9 as a function of probe and relation type. As in each of the preceding conditions, more presented than inferred information was recalled from both true-probed ($F(1,70)=97.01$, $p<.001$) and false-probed passages ($F(1,70)=108.31$, $p<.001$). As in condition 1, the main effects due to relation type were not obtained, but did interact with the distance variable (true probes: $F(1,70)=6.72$, $p<.025$; false probes: $F(1,70)=5.83$, $p<.025$). Recall of SI presented relations exceeded LO while LO inferences were more frequently recalled than SI inferences. These differences were all significant except in the case of false-probed, presented relation recall (true probes, presented: $t_{78}=3.85$, $p<.001$; false probes, presented: $t_{78}=4.01$, $p<.001$; false probes, inferred: $t_{78}=4.52$, $p<.001$). Consideration of recall associated only with errorless true probe performance did not change the pattern.

## Recognition

Scoring. Recognition data were collected from all eighty Ss in condition 3. The scores were converted into values on the ten point scales described in condition 1.

Figure 9

Condition 3: Mean Percentage of Presented and Inferred Relations Recalled from True- and False-Probed Passages

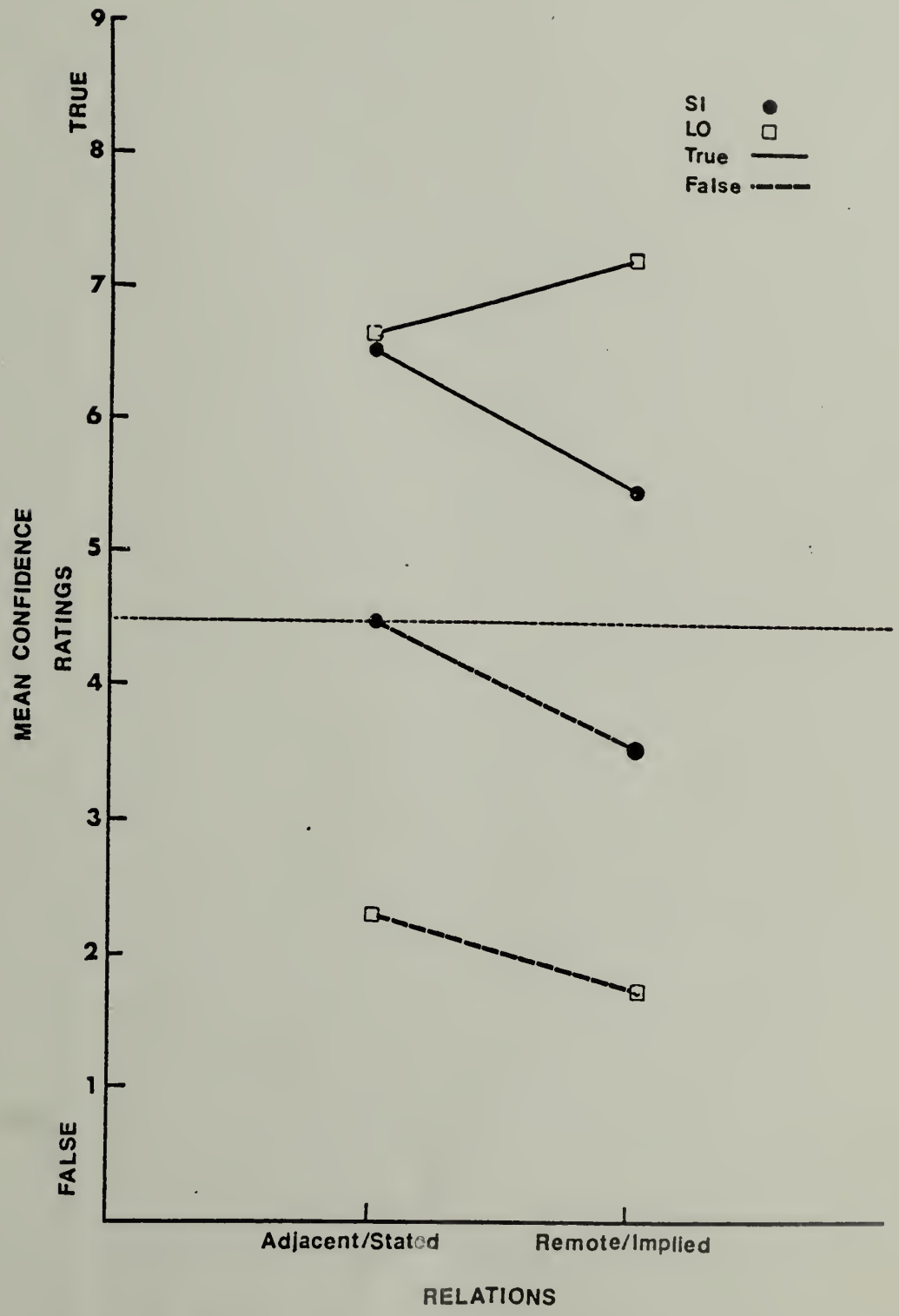


Recognition of True and False Items as a Function of Relation-type and Distance. The relevant data are summarized in Figure 10. They represent a replication of condition 1 as the probing manipulation did not influence recognition. The same overall patterns were obtained in both conditions as substantiated by a set of 16 contrasts. The experimentwise error rate was held at $\alpha$=.05 with critical $t_{39}$=2.86 and $t_{78}$=2.81. The presentation order of the adjacent relations within the passages allowed their relatively easy integration. The SDE was again obtained for LO. Ss ' performance was more accurate on both true and false remote LO items than adjacent (true: $t_{39}$=4.82; false: $t_{39}$=4.74). The SI means again resulted in a truth by distance interaction: accuracy decreased from adjacents to remotes for true items ($t_{39}$=5.12) and increased for false items ($t_{39}$=4.17).

Overall recall was significantly better in condition 3 than 1, so too was recognition (cf. Figure 2). All of the means for true items fell well within the true range (SI adjacent, $t_{39}$=7.61; SI remote, $t_{39}$=4.08; LO adjacent, $t_{39}$= 7.81; LO remote, $t_{39}$ =9.11). Ss were equally well able to recognize the adjacent relations which were actually presented regardless of relation type ($t_{78}$<1). LO true inferences were recognized as such with significantly more confidence than SI inferences ($t_{78}$=3.80). Once more performance was least accurate for SI false adjacent items.

Figure 10

Condition 3: Mean True-False Confidence Ratings

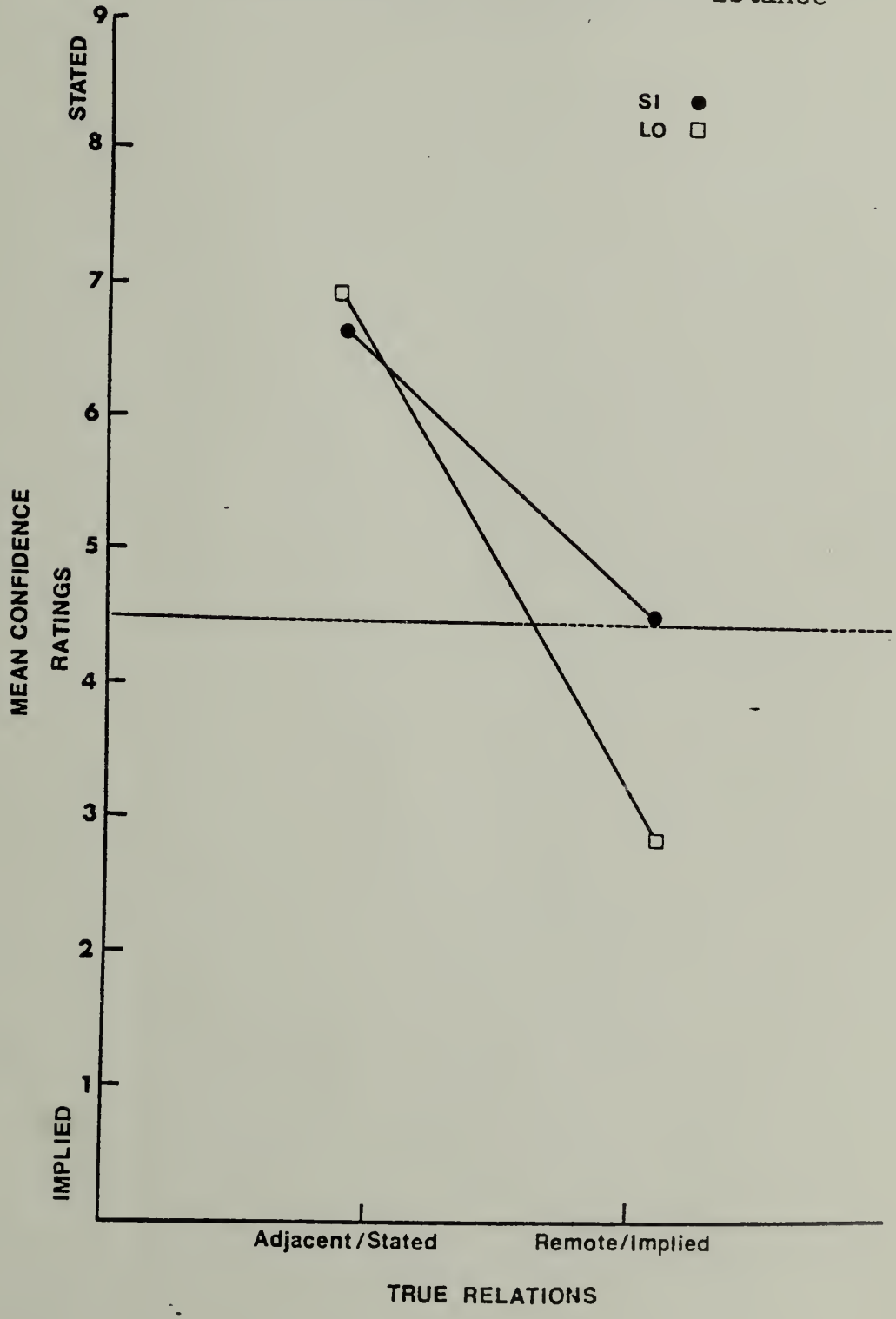


Their mean fell right at the guessing level of 4.5. The false SI inferences though were significantly rated within the false range ( $\bar{X} < 4.5$, $t_{39}=3.11$). Ss showed much more confidence in judging false LO relations as false than SI (adjacents, $t_{78}=5.52$; remotes, $t_{78}=5.94$).

Stated-Implied Judgments. The mean stated-implied ratings for true items in condition 3 can be seen in Figure 11. A main effect due to distance and its significant interaction with relation type were obtained ($F(1,78)=24.16$, $p<.001$, $F(1,78)=8.54$, $p<.005$, respectively). As in the other conditions, LO and SI were equally able to identify the relations they had actually heard ($t_{78}<1$). Given that a true inference was recognized as such, only in the case of LO were Ss really successful in acknowledging that the relation had been inferred. The LO remote mean was significantly within the implied range ($\bar{X}<4.5$, $t_{39}=6.21$, $p<.001$), while the SI mean did not differ from chance ($\bar{X}<4.5$, $t_{39}<1$). The same pattern was manifest in condition 1 though the interaction did not quite reach significance.

Figure 11

Condition 3: Mean Stated-Implied Confidence Ratings for True Items as a Function of Distance



## Discussion

### Basic Patterns in the Results

Probes. The true probe error rates obtained by Sefkow (1976) were due to faulty logic specific to SI as significantly fewer errors were made with identically constructed LO materials. As the size of the probe inferences increased, error rates also increased for SI as predicted by Ss' tendency to reject SI transitive inferences. At the same time error rates dropped for LO consistent with the SDE obtained with integrated LO. Interfering with the ease of integration did not affect SI true probe error rates, indicating that integration was not taking place at the time of initial encoding. The nonadjacency of the mediating SI relations was not critical to later location or integration of the links. However, hindering initial integration caused a significant increase in LO true probe errors suggesting that, unlike with SI, when possible LO Ss referred to an integrated representation in responding to the probes.

Recall. The BRE obtained by Sefkow (1976) was replicated with new SI materials. The verification of two-link inferences resulted in enhanced recall of the mediating links, independent of the ease of integration. When the LO materials were easily integrated, this pattern

of recall was not obtained though a diffuse overall advantage did accrue. Only when the initial integration was made difficult did facilitation of the mediating links occur. Thus, SI Ss appeared to be integrating the relevant links only at the time of the probe. In contrast, the data suggest that LO Ss had the inferential information available prior to the probe when the presentation order allowed it. In condition 2, LO Ss were forced to rely on memory for the original propositions as were the SI Ss; therefore, similar patterns of facilitation were obtained with the two relation types.

The results of condition 3 support the notion that only LO tend to be fully schematized when initially encoded. Correct responses to more extensive inferential probes facilitated recall of SI mediating links, particularly those not sharing an end term with a probe. The easily integrated LO materials enjoyed no recall advantage as a function of the true probes.

Recognition. The recognition data served to substantiate the patterns obtained with the true probes. In conditions 1 and 3 the SDE, indicating an integrated representation, occurred only in conjunction with LO. The difficult ordering in condition 2 did disrupt the integration process. No SDE was obtained. The SI data showed the usual truth by distance interaction; that is, Ss tended to assume

symmetry of the relations and rejected transitive inferences.

Nature of the Representations

The results of conditions 1 and 2 directly contradict Potts' (1976) proposal that SI and LO share the same form of integrated, memorial representation and are processed the same once logical errors are eliminated. Instead, the findings suggest that, whereas under favorable circumstances Ss do readily draw and store LO inferential information, a proposition based store is a more appropriate formulation for SI. Ease of integration as varied from condition 1 to 2 had little effect on SI true probe error rates, recognition, or recall patterns. Of particular interest was the replication of Sefkow's (1976) earlier finding in all three conditions: correctly verifying a true inference involving from two to four mediating SI relations resulted in later enhanced recall of those same relations, regardless of integrability. The true inference probes also enhanced LO recall in conditions 1 and 2; however, the SI pattern of facilitation was obtained for LO only in condition 2. The LO true probe error rates and the recognition data support the conclusion that Ss were integrating the LO relations in condition 1 but were much less successful in doing so in condition 2, presumably where they were forced to rely more heavily on memory for the original propositions.

Subordinate findings support the notion of a proposition

based representation for SI. Comparison of the recall of presented and inferred relations reflect the patterns expected from the two types of storage. Whether integration was difficult or not, Ss recalled the presented relations better for SI than LO. The generation and later recall of inferences was better for LO than SI only when integration was likely. When the difficult order was imposed, not only did LO inference recall fall below that of SI but recall of the original relations also fell sharply.

The latter finding would suggest that the integration of newly encountered LO propositions may be required not only to more meaningfully elaborate the interrelationships but for reasonable recall of the original relations. This was not the case for SI. One possible explanation which could account for these results is that single LO relations are highly confusable, particularly when the same comparative adjective relates each element to the next. The sharp increase in the recall of class elements outside the context of a relation and incorrect relation reversals (e.g.,BA) exclusively for LO in condition 2 supports this proposal (see Table 18). At the same time it may be the explicit focus on unidimensionality that makes the integration strategy so apparent and relatively easy to implement. In contrast, SI specify not one but many implicitly shared semantic attributes between classes or elements. Rather than concentrate on the general properties of SI, and

elaborate between relations as is done with LO, Ss may attend to some subset of the shared attributes and thus more fully elaborate within a relation. As Ss process each succeeding SI adjacent relation there is no guarantee of consistency in the selection of attributes which would be necessary for more general properties like transitivity to be made apparent. Thus, while the multidimensional aspect may make the individual relations more discriminable and less susceptible to interference, it may also hamper integration. This distinction between LO and SI needs to be more fully specified and tested empirically; however, at least one piece of evidence reported by Potts (1976) does lend some support to the notion. He found that Ss readily processed SI as LO (i.e., obtained the SDE) when the related items were nonsense syllables. An encoding strategy which focuses on implicitly shared attributes was made impracticable as no set of semantic features could be specified in any natural way. It can then be argued that, as with LO, Ss were forced to attend to what now can be viewed as the single salient attribute: "is a" and its general properties.

A number of subordinate findings replicate earlier results and help to further specify the general nature of the representations. For instance, a memory deficit was not the sole source of true probe errors as initially proposed by Sefkow. The pattern of such errors across

conditions, as well as the recognition data, reaffirmed that Ss tend to make logical errors specific to SI. That is, they erroneously assume symmetry of SI relations and reject legitimate transitive inferences. Unfortunately, available data do not allow the source of these logical errors to be identified. The present study does indicate; however, that difficulty in the initial encoding of SI relations as compared to LO cannot be claimed because recall of the presented information was consistently superior for SI while recognition levels were equivalent.

The stated-implied ratings indicated that for both relation types, in all three conditions, Ss could distinguish between what was presented and what was inferred, though their confidence in the former was stronger. This result adds to a growing literature (e.g., Lawson, 1977) which argues against the classical Bransford and Franks (1971) conclusion that only the integrated, wholistic representation is retained in memory.

Nature of the Backward Review Effect

Once again, it was demonstrated that correctly answering a question directed at information available only in memory facilitates retention of that material. The exact nature of the facilitative effect depended upon the probe requirements, as well as the memorial representation of the passage

information. In the case of SI, the effect was concentrated on those relations whose integration was necessary for the probes' verification. This pattern did not change as a function of ease of integration; hence, the BRE obtained with SI was not due to strengthening a previously integrated representation.

Performance on LO passages mirrored that of SI only when initial integration of the links was hampered via the difficult presentation order of condition 2. When Ss listened to the more easily integrated passages, different recall patterns emerged. Two-link probes resulted in a nonspecific increase in both link and inference recall. This finding can be adapted to at least some current models which try to account for the integration of LO and the resulting SDE. For example, a spreading activation notion incorporated into a Potts-type rating scale model (1974) could predict this result.

If all mediating SI relations are facilitated by an inferential probe, then the use of more extensive probes should result in better link recall. Facilitation of SI relations in condition 3 was not limited to the mediating links nor was it as strong as expected. Several points need to be made. First, the true probe error rate was extremely high which reduced the size of the sample when corrects only were considered. Second, some Ss do recognize that SI can be elaborated just as LO and readily

do so (Potts, 1976). It is likely that as the span of the true probes (number of mediating relations) increases, a larger percentage of those responding correctly are doing so on the basis of an integrated representation. In these cases, the usual SI pattern of facilitation would not be expected.

On the other hand, three and four link probes did not facilitate LO recall relative to performance on passages subject to false irrelevant probes. A major unresolved finding is that the latter was surprisingly good when compared to the identically probed passages of condition 1. A forward orienting effect due to differences in true probe difficulty was unsupported. No good theoretical account of the data seems readily apparent. Hence, it appears likely that the differences were due to S variability and/or some type of ceiling effects.

In condition 2, the order of the SI relations within the passages did appear to lessen the magnitude of the BRE (component minus irrelevant link recall : 26.1% versus 17.0%). If the drop was due to a difficulty in locating and/or integrating the relevant, component relations, one would expect an attendant increase in true probe errors. On the contrary, slightly fewer errors were made. The exact role that presentation order plays in the enhancement of SI relations needs to be more closely examined.

The order manipulation provided one additional piece

of information regarding the nature of backward review effects. Enhanced recall of the previously unintegrated SI and LO propositions was not due to some simple ordered scan of memory. The recall of irrelevant links which were physically within the span of the probes was not affected. A process involving direct access of the relevant propositions could well account for this finding.

The present study does not provide evidence as to whether the BRE obtained with SI and unintegrated LO materials was basically due to the rehearsal and/or integration of the propositional units. This question remains open for further research as both are viable alternatives.

References

Anderson, R.C. & Biddle, W.B. On asking people questions about what they are reading. In G. Bower (Ed.), The Psychology of Learning and Motivation, Vol. 9, New York: Academic Press, 1976.

Anderson, J.R. & Bower, G.H. Human Associative Memory. Washington D.C.: Winston, 1973.

Barclay, J.R. The role of comprehension in remembering sentences. Cognitive Psychology, 1973, 4, 229-254.

Bartlett, F.C. Remembering: A Study in Experimental Psychology. Cambridge: Cambridge University Press, 1932.

Bransford, J.D., Barclay, J.R. & Franks, J.J. Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 3, 193-209.

Bransford, J.D. & Franks, J.J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Bransford, J.D. & Franks, J.J. The abstraction of linguistic ideas: A review. Cognition: International Journal of Cognitive Psychology, 1972, 1, 211-249.

Bransford, J.D. & Johnson, M.K. Consideration of some problems of comprehension. In W. Chase (Ed.), Visual Information Processing, New York: Academic Press, 1973.

Buckley, P.B. & Gillman, C.B. Comparisons of digits and dot patterns. Journal of Experimental Psychology, 1974, 103, 1131-1136.

Carroll, M. & Kammann, R. The dependency of schema formation

on type of verbal material: Linear orderings and set inclusions. Memory and Cognition, 1977, 5, 73-78.

Ceraso, J. & Provitera, A. Sources of error in syllogistic reasoning. Cognitive Psychology, 1971, 2, 400-410.

Cofer, C.N. Constructive processes in memory. American Scientist, 1973, 61, 537-543.

Craik, F.M. & Tulving, E. Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268-294.

Foos, P.W., Smith, K.H., Sabol, M.A. & Mynatt, B.T. Constructive processes in simple linear-order problems. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 759-766.

Frase, L.T. Structural analysis of the knowledge that results from thinking about a text. Journal of Educational Psychology Monograph Supplement, 1969, 60, 1-16.

Frase, L.T. Inference of sentence order and amount of higher level text processing upon reproductive and productive memory. American Educational Research Journal, 1970, 7, 307-319.

Griggs, R.A. Logical processing of set inclusion relations in meaningful text. Memory and Cognition, 1976, 4, 730-740.

Holyoak, K.J. & Walker, J.H. Subjective magnitude information in semantic orderings. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 287-299.

Humphreys, M.S. The derivation of endpoint and distance effects in linear orderings from frequency information. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 496-505.

Huttenlocher, J. Constructing spatial images: A strategy in reasoning. Psychological Review, 1968, 75, 550-560.

Jenkins, J.J. Remember that old theory of memory? Well forget it! American Psychologist, 1974, 29, 785-795.

Johnson, D.M. Systematic Introduction to the Psychology of Thinking. New York: Harper & Row, 1972.

Lawson, R. Representation of individual sentences and holistic ideas. Journal of Experimental Psychology: Human Learning and Memory, 1977,3, 1-9.

McGaw, B. & Grotelueschen, A. Directions of the effect of questions in prose materials. Journal of Educational Psychology, 1972, 63, 580-588.

McKinley, J.E. Utilizing comparative information from memory: The symbolic distance effect. Unpublished Master's Thesis, University of Massachusetts, Amherst, 1976.

Moyer, R.S. Comparing objects in memory: Evidence suggesting an internal psychophysics. Perception and Psychophysics, 1973, 13, 180-184.

Moyer, R.S. & Bayer, R.H. Mental comparisons and the symbolic distance effect. Cognitive Psychology, 1976, 8, 228-246.

Moyer, R.S. & Landauer, T.K. Time required for judgements of numerical inequality. Nature (London), 1967, 215,1519-20.

Paivio, A. Perceptual comparisons through the mind's eye. Memory and Cognition, 1975, 3, 635-647.

Parkman, J.M. Temporal aspects of digit and letter inequality judgements. Journal of Experimental Psychology, 1971, 91, 191-205.

Potts, G.R. Information processing strategies used in the encoding of linear orderings. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 727-740.

Potts, G.R. Incorporating quantitative information into a linear ordering. Memory and Cognition, 1974, 2, 533-538.

Potts, G.R. Storing and retrieving information about ordered relationships. Journal of Experimental Psychology, 1974, 103, 431-439.

Potts, G.R. Artificial logical relations and their relevance to semantic memory. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 746-758.

Quillian, M.R. The teachable language comprehender: A simulation program and theory of language. Communications of the ACM, 1969, 12, 459-476.

Riley, C.A. & Trabasso, T. Comparatives, logical structures, and encoding on a transitive inference task. Journal of Experimental Child Psychology, 1974, 102, 323-326.

Rips, L.J., Shoben, E.J. & Smith, E.E. Semantic distance and verification of semantic relations. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 1-20.

Rothkopf, E.Z. Learning from written materials: An exploration

of the control of inspection behaviors by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Rothkopf, E.Z. & Billington, M.J. Indirect review effects produced by adjunct questions: A replication study. Journal of Educational Psychology, 1974, 66, 669-679.

Scholz, K.A. & Potts, G.R. Cognitive processing of linear orderings. Journal of Experimental Psychology, 1974, 102, 323-326.

Sefkow, S.B. Review effects of adjunct questions on learning from prose (Report No. 76-2). Amherst, Massachusetts: University of Massachusetts, Department of Psychology, Cognitive Processes Laboratory.

Smith, E.E., Shoben, E.J. & Rips, L.J. Structure and process in semantic memory: A feature model for semantic decision. Psychological Review, 1974, 81, 214-241.

Trabasso, T., Riley, C.A. & Wilson, E.G. The representation of linear order and spatial strategies in reasoning: A developmental study. In R. Falmagne (Ed.), Reasoning: Representation and Process, New York: LEA, 1975.

Tzeng, O.J. Sentence memory: Recognition and inference. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 720-726.

Watts, G.H. & Anderson, R.C. Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 1971, 62, 387-394.

# Appendix A

Passage 1: Set Inclusion, Adjacent Order

During orientation week, some freshmen at Potter College were amazed by the great care taken to explain the intricacies of the library system. For example, they were informed that all the books shelved in the North Annex had been contributed by the Lilly Foundation, a philanthropic organization. All oversized books are located in the North Annex because of their special status. All books with specially reinforced bindings are necessarily oversized. Since there is much demand for them, all reference books have specially reinforced bindings.

Passage 1: Set Inclusion, Nonadjacent Order

During orientation week, some freshmen at Potter College were amazed at the great care taken to explain the intricacies of the library system. For example, all oversized books are located in the North Annex because of their special status. Since there is much demand for them, all reference books have specially reinforced bindings. They were also informed that all books shelved in the North Annex had been contributed by the Lilly Foundation, a philanthropic organization. Furthermore, all books with specially reinforced bindings are necessarily oversized.

Passage 1: Linear Ordering, Adjacent Order

During orientation week, some freshmen at Potter College were amazed by the great care taken to explain the intricacies of the library system. For example, they were informed that all the books shelved in the North Annex were more expensive than those contributed by the Lilly Foundation, a philanthropic organization. All oversized books are more expensive than those located in the North Annex, because of their special status. All books with specially reinforced bindings are necessarily more expensive than those that are oversized. Since there is much demand for them, all reference books are more expensive than those with specially reinforced bindings.

Passage 1: Linear Ordering, Nonadjacent Order

During orientation week, some freshmen at Potter College were amazed at the great care taken to explain the intricacies of the library system. For example, all oversized books are more expensive than those located in the North Annex because of their special status. Since there is much demand for them, all reference books are more expensive than those with specially reinforced bindings. They were also informed that all books shelved in the North Annex were more expensive than those contributed by the Lilly Foundation, a philanthropic organization. Furthermore, all books with specially reinforced bindings are necessarily more expensive than those that are oversized.

Passage 2: Set Inclusion, Adjacent Order

In the March 1974 issue of Sportsmen Magazine, James Freeband, the noted fisherman, wrote in his column that all snail eating fish are susceptible to dorsal fin disorders. He also noted that, in his experience, he had found that all cold water fish eat snails. In the previous month's column, Mr. Freeband wrote of an important discovery by Thomas Gibbs that all endangered species of fish are cold water fish. It is common knowledge among fishermen that Mr. Gibbs is very reputable. So when he states that the Dogel fish is on the endangered species list, we can take him seriously.

Passage 2: Set Inclusion, Nonadjacent Order

In the March 1974 issue of Sportsmen Magazine, James Freeband, the noted fisherman, wrote in his column that all cold water fish eat snails. He also noted that the Dogel fish is on the endangered species list. In the previous month's column, Mr. Freeband wrote of an important discovery by Thomas Gibbs that all snail eating fish are susceptible to dorsal fin disorders. It is common knowledge among fishermen that Mr. Gibbs is very reputable. So when he states that all endangered species of fish live in cold water, we can take him seriously.

Passage 2: Linear Ordering, Adjacent Order

In the March 1974 issue of Sportsmen Magazine, James Freeband, the noted fisherman, wrote in his column that all snail eating fish are more aggressive than those susceptible to dorsal fin disorders. He also noted that, in his experience, he had found that all cold water fish are more aggressive than those that eat snails. In the previous month's column, Mr. Freeband wrote of an important discovery by Thomas Gibbs that all endangered species of fish are more aggressive than cold water fish. It is common knowledge that Mr. Gibbs is very reputable. So when he states that the Dogel fish is more aggressive than those on the endangered species list, we can take him seriously.

Passage 2: Linear Ordering, Nonadjacent Order

In the March 1974 issue of Sportsmen Magazine, James Freeband, the noted fisherman, wrote in his column that all cold water fish are more aggressive than fish that eat snails. He also noted that the Dogel fish is more aggressive than fish on the endangered species list. In the previous month's column, Mr. Freeband wrote of an important discovery by Thomas Gibbs that all snail eating fish are more aggressive than those fish susceptible to dorsal fin disorders. It is common knowledge among fishermen that Mr. Gibbs is very reputable. So when he states that all endangered species of fish are more aggressive than cold water fish, we can take him seriously.

Passage 3: Set Inclusion, Adjacent Order

For many years, anthropologists have been particularly interested in studying the culture and values of the primitive Wambi Tribe. The Wambis are known for their complex caste-like system. For example, all tatooed men are priests. According to a custom dating back at least 500 years, each warrior is tatooed. Traditionally, every Wambi farmer is a warrior. The possibility of long and serious wars is dreaded by the tribe. At an early age all Wambi male children are trained in the techniques of farming, as the future prosperity of the tribe rests with them.

Passage 3: Set Inclusion, Nonadjacent Order

For many years, anthropologists have been particularly interested in studying the culture and values of the primitive Wambi Tribe. The Wambis are known for their complex caste-like system. For example, according to a custom dating back at least 500 years, each warrior is tatooed. At an early age, all Wambi male children are taught the techniques of farming, as the future prosperity of the tribe rests with them. All tatooed Wambi men become priests. Traditionally every Wambi farmer is also trained to be a warrior. The possibility of long and serious wars is dreaded by the tribe.

Passage 3: Linear Ordering, Adjacent Order

For many years, anthropologists have been particularly interested in studying the culture and values of the primitive Wambi Tribe. The Wambis are known for their complex caste-like system. For example, all tatooed men are more honored than priests. According to a custom dating back at least 500 years, all warriors are more honored than tatooed men. Traditionally, all Wambi farmers are honored more than the warriors. The possibility of long and serious wars is dreaded by the tribe. At an early age all Wambi male children are honored more than the farmers, as the future prosperity of the tribe rests with them.

Passage 3: Linear Ordering, Nonadjacent Order

For many years, anthropologists have been particularly interested in studying the culture and values of the primitive Wambi Tribe. The Wambis are known for their complex caste-like system. For example, according to a custom dating back at least 500 years, all warriors are more honored than the tatooed men. At an early age, all Wambi male children are more honored than the farmers, as the future prosperity of the tribe rests with them. All tatooed Wambi men are more honored than the priests. Traditionally, all Wambi farmers are more honored than the warriors. The possibility of long and serious wars is dreaded by the tribe.

Passage 4: Set Inclusion, Adjacent Order

Many gardeners have found the services of the Ramsey County Agricultural Service to be very useful. Each spring they make available pamphlets containing the latest gardening developments. One recent publication stated that all flowers blooming in April must be exposed to full sun. This fact was previously unknown. Furthermore, all plants set in the ground in October will bloom in April. The planting information also asserted that tuberous plants must all be set in the ground in October. For those interested in special garden layouts, the availability of a new hybrid was announced, the Gloxolia. It is a new tuberous variety.

Passage 4: Set Inclusion, Nonadjacent Order

Many gardeners have found the services of the Ramsey County Agricultural Service to be very useful. Each spring they make available pamphlets containing the latest gardening developments. One recent publication stated that all plants set in the ground in October will bloom in April. For those interested in special garden layouts, the availability of a new hrbrid was announced, the Gloxolia. It is a new tuberous variety. In addition, it was stated that all flowers that bloom in April must be exposed to full sun. This fact was previously unknown. The planting information also asserted that all tuberous plants must be set in the ground in October.

Passage 4: Linear Ordering, Adjacent Order

Many gardeners have found the services of the Ramsey County Agricultural Service to be very useful. Each spring they make available pamphlets containing the latest gardening developments. One recent publication stated that all flowers blooming in April are more colorful than flowers exposed to full sun. Furthermore, all plants set in the ground in October are more colorful than those that bloom in April. The planting information also asserted that all tuberous plants are more colorful than plants set in the ground in October. For those interested in special garden layouts, the availability of a new hybrid was announced, the Gloxolia. It is more colorful than tuberous plants.

Passage 4: Linear Ordering, Nonadjacent Order

Many gardeners have found the services of the Ramsey County Agricultural Service to be very useful. Each spring they make available pamphlets containing the latest gardening developments. One recent publication stated that all plants set in the ground in October are more colorful than those that bloom in April. For those interested in special garden layouts, the availability of a new hybrid was announced, the Gloxolia. It is more colorful than tuberous plants. In addition, it was stated that all flowers that bloom in April are more colorful than flowers exposed to full sun. The planting information also asserted that all tuberous plants are more colorful than plants set in the ground in October.

Passage 5: Set Inclusion, Adjacent Order

Mr. Crane, an American food specialist, is being sent to South America to consult with government officials. First, he will have to thoroughly familiarize himself with the characteristics of the foods produced there, as well as their government policies. For example, all fruit which has been stored in special oxygen chambers is exported. Because of local conditions, all apples grown in the mountains are stored in oxygen chambers. The climate and soil are such that all Peruvian fruit is grown in the mountains. The new seedless apples are grown in Peru.

Passage 5: Set Inclusion, Nonadjacent Order

Mr. Crane, an American food specialist, is being sent to South America to consult with government officials. First, he will have to thoroughly familiarize himself with the characteristics of the foods produced there, as well as their government policies. For example, because of local conditions all apples grown in the mountains must be stored in special oxygen chambers. The new seedless apples are all grown in Peru. All fruits which have been stored in special oxygen chambers must be exported for sale. The climate and soil are such that all Peruvian fruit is grown in the mountains.

Passage 5: Linear Ordering, Adjacent Order

Mr. Crane, an American food specialist, is being sent to South America to consult with government officials. First, he will have to thoroughly familiarize himself with the characteristics of the foods produced there, as well as their government policies. For example, all fruit which has been stored in special oxygen chambers is sweeter than exported fruit. Because of local conditions, all apples grown in the mountains are sweeter than those stored in special oxygen chambers. The climate and soil are such that all Peruvian fruit is sweeter than mountain grown fruit. The new seedless apples are all sweeter than apples grown in Peru.

Passage 5: Linear Ordering, Nonadjacent Order

Mr. Crane, an American food specialist, is being sent to South America to consult with government officials. First, he will have to thoroughly familiarize himself with the characteristics of the foods produced there, as well as their government policies. For example, because of local conditions all apples grown in the mountains are sweeter than those stored in special oxygen chambers. The new seedless apples are all sweeter than those grown in Peru. All fruits which have been stored in special oxygen chambers are sweeter than those that must be exported for sale. The climate and soil are such that all Peruvian fruit is sweeter than fruit grown in the mountains.

# Appendix B

## Passage Instructions

You are going to listen to a recording of five short ficticious passages, each on a different topic. After each passage you will be given fifteen seconds to judge a statement true or false based solely on the content of the passage you just heard. Once you have verified whether it is valid or not, enter the appropriate answer, true or false, on the answer sheet provided. If you have no idea whether the statement is valid or not, do not guess; simply put a question mark in the answer space. The statements are in the booklet face down in front of you. Following the first passage you will hear a click. Turn the booklet over and respond to the first question. After the fifteen second answer period is up, you will hear a second click followed by the next passage. Remember, do not turn the booklet page to the next question until you hear the click following the appropriate passage. After you have completed this task, you will be further tested on what you have learned. Are there any questions? We will now begin.

## Recall Instructions

Now I would like you to write down everything that you learned from the five passages you have just heard. You will no doubt find verbatim recall difficult; just be sure

to write down anything and everything that you can recall about each passage. I will tell you when to begin. There is a page for each passage in the booklet before you. If you need more room, use the back of the sheet. The topics for each passage are given at the head of each page. Do them in the order in which they are presented, and do not turn back once you have completed a topic. You will have $3\frac{1}{2}$ minutes to recall each passage. I will signal you when you are to begin and when to start each additional passage. Are there any questions? Begin.

Recognition Instructions

One last task is involved, a recognition test. A number of sentences are written in the booklet before you. Your task is to first decide whether each sentence is true or false based on the paragraphs you heard earlier. Second, you are to rate how confident you are of that answer on a scale from one to five where one means very low confidence and five means very high confidence. Third, if you decided the sentence was true, you must decide if it was explicitly stated or merely implied and again rate your confidence in this answer from one to five.

Circle your responses on the answer sheet. Please use the full range of confidence ratings. You are to work through the booklet at your own speed. Be sure not to look back once you have turned a page. Are there any questions? Please begin.